APPLICATION OF LINEAR STOCHASTIC OPERATOR THEORY

By L. H. Sibul

Technical Memorandum
File No. TM 662.2741-05
October 21, 1968
Contract NOw 65-0123-d
Copy No. 5

The Pennsylvania State University
Institute for Science and Engineering
ORDNANCE RESEARCH LABORATORY
University Park, Pennsylvania

NAVY DEPARTMENT · NAVAL ORDNANCE SYSTEMS COMMAND

17

ABSTRACT

In diverse areas of physics and engineering, problems arise which should properly be described by linear differential equations with stochastic coefficients. Methods are developed here for finding integral expressions for the second-order statistics (means, correlation functions and power spectrum) of the dependent variable of an n^{th} order linear stochastic differential equation. These expressions constitute a generalization of the corresponding expressions for linear time-varying systems to linear randomly time-varying systems. The kernels of the integral expressions for the statistical measures of the solution can be interpreted as stochastic Green's functions.

In general, expressions for the second-order statistics of the solution of either ordinary or partial linear differential equations with stochastic coefficients requires knowledge of all the moments of stochastic coefficients. An exceptional case is that in which the stochastic coefficients are Gaussian processes. Then the knowledge of the second-order statistics is sufficient for the complete solution.

It is assumed that the coefficients of the differential equation are separable into deterministic and stochastic parts, and the solution for the deterministic part is known. In the case of a stochastic ordinary differential equation, the problem now becomes a problem of solving a Volterra integral equation with a stochastic kernel. Two methods of solution of this integral equation are considered: the Neumann series expansion method and the degenerate kernel method. A theorem which gives sufficient conditions for the uniform convergence of the Neumann series expansion is proved. The proof of this theorem, and the actual Neumann series expansion, is shown to be facilitated

if the n^{th} order differential equation is expressed in the state-space notation of control system theory.

The uniform convergence of the Neumann series expansion allows the solution of the stochastic differential equation to be expressed in terms of the resolvent kernel of the stochastic integral equation. The ensemble average and the covariance function of the solution are expressed in terms of the corresponding statistical measures of the resolvent kernel and of the input process. The statistical measures of the resolvent kernel are functions of both the Green's function of the deterministic operator and the appropriate statistical measures of the stochastic coefficients.

Both the Neumann series iteration and degenerate kernel approximation are applied to the investigation of the propagation of waves in a randomly space- and time-varying medium. Almost all the previous work has used the so-called quasimonochromatic assumption which essentially neglects the time variation of the medium. Such an assumption has been avoided in this dissertation and thereby some consequences of this assumption discovered. The source and the stochastic medium are assumed to be wide-sense stationary stochastic processes. All the stochastic quantities of the scalar wave equation are expressed by their spectral representation, and the equation is solved for the spectral representation of the scalar wave function. From the spectral representation of the scalar wave function, its power spectral density and mutual coherence functions can be found. Both the Neumann series expansion and the degenerate kernel approximation demonstrate the spreading of the power spectrum of the source by the time-varying

medium. In the Neumann series expansion, even the first-order approximation shows the spreading of the power spectrum. Higher order approximations show further spreading of the power spectrum. Higher order approximations also show that the solution contains, in addition to the wide-sense stationary terms, terms which are no longer wide-sense stationary.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

1.1 General Statement of the Problem

Many problems of interest to electrical engineers are described by linear differential equations with stochastic* coefficients. Often the randomness of the coefficients has been neglected because no widely applicable and tractable mathematical methods have been known for solving such problems. In this dissertation, an n^{th} order linear differential equation with such stochastic coefficients

$$\mathcal{L}\, y(t,\omega) = x(t,\omega) \tag{1.1}$$

is considered. $\mathcal{L}$ is assumed to be a sum of an invertible deterministic operator L and a stochastic operator R. The differential operator $\mathcal{L}$ is defined on some domain $t\epsilon T$ and a probability space $(\Omega,\mathcal{X},P)$.** L can be either an ordinary or a partial differential operator. Our approach to the solution of (1.1) is to determine a "stochastic Green's function" (Adomian, 1961, 1963, 1964) for the linear stochastic operator $\mathcal{L}$ in terms of a deterministic Green's function for L and the appropriate "statistical measures" of the stochastic coefficients. The term "statistical measures" is used as a general term for the quantities that characterize stochastic processes. For example,

---

* Random functions. Precise definitions are given in Chapter II, section 2.2.

** Probability spaces are also defined in section 2.2.

expectations, or averages, spectral densities and correlation functions are statistical measures. The integral kernel which expresses the desired statistical measure of an output process in terms of the corresponding statistical measure of an input and appropriate statistical measures of the stochastic coefficients is called the "stochastic Green's function". In this thesis, an iterative method (Adomian, 1967) for the construction of stochastic Green's functions is investigated and applied to physical problems.

1.2 Significance of the Problem

Differential equations with randomly time-varying coefficients arise naturally in many practical problems. Analysis of seemingly simple systems often produces such differential equations. For example, to find the current drawn from a generator with finite internal impedance operating into a randomly time-varying load impedance one must solve a differential equation with stochastic coefficients. A large interconnected power system is clearly a randomly time-varying network. Its behavior and instantaneous states are only predictable in a statistical sense. Many control system problems have randomly time-varying parameters. Adaptive control systems are typical examples. In other cases, system parameters are modulated by random disturbances. The combination of complexity and uncertainty in a real problem often makes it necessary to use stochastic analysis. Uncertainty may result from unavoidable experimental errors in determining the parameters of the real system or it may result from lack of prior knowledge of the conditions under which a system must operate. In the latter case, statistical analysis

is preferable to the "worst case analysis", because the results of the worst case analysis are often unrealistically pessimistic.

In some cases it is preferable to analyze a complex deterministic system as if it were a simpler stochastic system. For example a many-body problem may be attacked by considering the interaction of the predominating forces as deterministic forces and the totality of all smaller forces as random forces. Note that the sum of all the small forces may not be negligible compared to the larger forces, even if each individual force by itself is negligible. In electrical engineering, there are analogous problems, such as complicated network problems which would require simultaneous solution of a large number of loop or node equations. Here, also, the complexity can be traded for randomness.

In addition to the above examples, another important class of problems involving stochastic differential equations arises in almost all wave propagation problems. The propagation of electromagnetic waves through the atmosphere, troposphere, plasmas, turbulent mixtures of gases and water vapor are some examples. The propagation of sound waves through water with varying temperature gradients, microstructure or turbulence, results in stochastic scalar wave equations. Thus the stochastic wave equations are of interest to sonar, radar and communication engineers. Furthermore, closer understanding of stochastic wave equations may provide additional research tools for understanding various random media.

In addition to the above-mentioned examples stochastic operator equations are significant in several fundamental problems of physics. In spite of this, the stochastic operator approach has not been widely

expectations, or averages, spectral densities and correlation functions are statistical measures. The integral kernel which expresses the desired statistical measure of an output process in terms of the corresponding statistical measure of an input and appropriate statistical measures of the stochastic coefficients is called the "stochastic Green's function". In this thesis, an iterative method (Adomian, 1967) for the construction of stochastic Green's functions is investigated and applied to physical problems.

1.2 Significance of the Problem

Differential equations with randomly time-varying coefficients arise naturally in many practical problems. Analysis of seemingly simple systems often produces such differential equations. For example, to find the current drawn from a generator with finite internal impedance operating into a randomly time-varying load impedance one must solve a differential equation with stochastic coefficients. A large interconnected power system is clearly a randomly time-varying network. Its behavior and instantaneous states are only predictable in a statistical sense. Many control system problems have randomly time-varying parameters. Adaptive control systems are typical examples. In other cases, system parameters are modulated by random disturbances. The combination of complexity and uncertainty in a real problem often makes it necessary to use stochastic analysis. Uncertainty may result from unavoidable experimental errors in determining the parameters of the real system or it may result from lack of prior knowledge of the conditions under which a system must operate. In the latter case, statistical analysis

is preferable to the "worst case analysis", because the results of the worst case analysis are often unrealistically pessimistic.

In some cases it is preferable to analyze a complex deterministic system as if it were a simpler stochastic system. For example a many-body problem may be attacked by considering the interaction of the predominating forces as deterministic forces and the totality of all smaller forces as random forces. Note that the sum of all the small forces may not be negligible compared to the larger forces, even if each individual force by itself is negligible. In electrical engineering, there are analogous problems, such as complicated network problems which would require simultaneous solution of a large number of loop or node equations. Here, also, the complexity can be traded for randomness.

In addition to the above examples, another important class of problems involving stochastic differential equations arises in almost all wave propagation problems. The propagation of electromagnetic waves through the atmosphere, ionosphere, plasmas, turbulent mixtures of gases and water vapor are some examples. The propagation of sound waves through water with varying temperature gradients, microstructure or turbulence, results in stochastic scalar wave equations. Thus the stochastic wave equations are of interest to sonar, radar and communication engineers. Furthermore, proper understanding of stochastic wave equations may provide additional research tools for understanding various random media.

In addition to the abovementioned examples stochastic operator equations are significant in several fundamental problems of physics. In spite of this, the stochastic operator approach has not been widely

used because of mathematical difficulties. For this reason it is important to develop tractable mathematical methods for solving stochastic differential equations which arise in the problems of physics and engineering. According to the distinguished mathematician Solomon Lefschetz (1967), the study of stochastic differential equations is one of the fertile but poorly charted fields of mathematical research.

Formulation of physical problems by means of stochastic differential equations is a generalization of the corresponding formulation by means of deterministic differential equations. Conversely, one may consider all the coefficients of any differential equation to be random variables or functions; then, when the randomness of the coefficients vanishes, the problem reduces to the deterministic case, i.e., the random coefficients are replaced with their mean values. Problems which need to be investigated include the determination of whether randomness can be neglected in a given problem and what errors may result if this is done. A realistic approach to such an investigation would be to introduce the coefficient as a random variable, obtain the "random solution" of the differential equation, and then study its statistical properties. The general approach to this problem is discussed later in this chapter.

1.3 Previous Related Work

Before reviewing the previous related work in the theory of stochastic equations, some terminology needs to be clarified. The solution of a differential equation can be random for any of three reasons:

1) Initial or boundary conditions are random.

2) Forcing functions are stochastic processes.

3) The coefficients of the differential equation are stochastic processes.

In the first two cases, the differential equation itself is deterministic and the solution of the differential equation can be expressed in terms of a deterministic Green's function and random forcing functions, or random initial and boundary conditions. The random initial and boundary conditions may simply be interpreted as random forcing functions in the construction of the Green's function. For this reason, the first two cases are not really stochastic differential equations. Two stochastic processes are simply related by a deterministic Green's function. In this dissertation, the term "stochastic differential equation" means a differential equation with coefficients which are stochastic processes. Similarly, a stochastic integral equation is an integral equation with a stochastic kernel. In general, a "stochastic operator" is an operator with stochastic parameters.

For the first two cases, an excellent review of the solution of differential equations with random initial conditions and random forcing functions has been made by R. Syski (1967). Other useful survey papers are by Kampé de Férriet (1965) and an older one by Edwards and Moyal (1955). Syski (1967) also reviews other special topics such as Brownian motion problems or Wiener-Levy processes. Important mathematical contributions in connection with the Wiener-Levy processes have been made by Kolmogorov, Feller, Levy, Doob,

Kac, Itô and others. Syski also reviews the work that has been done by Itô and Doob in characterizing Markoff diffusion processes. The computation of the output spectral density and the correlation function for linear time-invariant systems has been thoroughly discussed in standard texts (Laning and Battin, 1956; Davenport and Root, 1958; and Papoulis, 1965) and needs no further comment.

Among the other works in this line, A. D. Jacobson (1967) in two recent papers analyzes second-order coherence properties of electromagnetic fields which are produced by random sources of arbitrary spectral width. His principal field quantity, "the dyadic field spectral density" is interpreted from both a statistical and a physical standpoint.

Also, random boundary conditions have been used by D. E. Barrick (1965) to study backscattering of electromagnetic waves from rough surfaces. The starting point of his analysis is the Chu-Stratton vector integral equation. He obtains closed form solutions for the backscattering cross sections for a class of rough surfaces with several different statistical models. Barrick's dissertation also contains an extensive list of references on backscattering of electromagnetic waves from rough surfaces. Another random boundary value problem which has been studied by several workers is the backscattering of sound waves from turbulent sea surface (Eckart, 1953; and Clay, 1960).

The study of linear differential equations with stochastic coefficients (or, in the more general case, linear stochastic operators) has proceeded along two lines. The first one, explicit solutions of specific problems of physics and engineering have been pursued by

Rosenbloom (1954), Tikhonov (1958), Samuels and Erdingen (1959), Aström (1965), Chelpanov (1962), Adomian (1961, 1963, 1964) and others. The most general approach to the subject has been taken by Adomian. The second approach has been based on probabilistic functional analysis and has been concerned with proving existence and uniqueness theorems. Work along this line has been done by Hanš (1961), Špaček (1955) and Bharucha-Reid (1960, 1964, 1965). As the theory of linear stochastic operators matures, it is expected that the two lines of research will merge.

Samuels and Eringen (1959) treated the problem of an n^{th} order linear differential equation with random coefficients. They restricted their attention to differential equations with (i) small randomly varying parameters, (ii) slowly varying random coefficients, and (iii) only one random coefficient. They applied their mathematical methods to an RLC circuit with a randomly varying capacitor and to the analysis of dynamic instability of an elastic bar subject to a randomly time-varying axial force. They used a perturbation method to solve the problem. Tikhonov (1958) has calculated the statistics of the solution of a first-order, linear, differential equation with a single stochastic coefficient and a stochastic forcing function. He assumed that the coefficient and forcing function were correlated, stationary and normally distributed. Tikhonov points out that sometimes it is possible to analyze the effects of a random forcing function on nonlinear systems by considering the solution of the abovementioned linear differential equation. Aström (1965) also considers a first order stochastic differential equation with correlated forcing function and coefficient. This problem arose in the

study of control systems subject to random distrubances. His analysis emphasized the probability distributions, where Tikhonov's work stressed correlation functions. Caughey and Dienes (1962) considered an n^{th} order linear differential equation with the forcing function and lowest order coefficient being white-noise processes. Syski (1967) reviews several other special differential equations with stochastic coefficients that have been solved. Other special problems have also been solved. For example, Gibson (1967) obtained a solution to a partial differential equation with random coefficients. The differential equation arose in calculation of interaction of electromagnetic waves and the random fluctuation of electron density in the wake of a reentry vehicle. The coefficients of the equation were assumed to be identically distributed Gaussian random variables.

In addition to the abovementioned methods, hierarchy techniques have been widely used by physicists for solution of differential equations with stochastic coefficients (Adomian, 1967; Adomian, 1968; Keller, 1964; Richardson, 1964; and Kraichnan, 1962). In the hierarchy methods, the differential equations are averaged before attempting to solve them. But by doing this, the equation for the average (first moment) involves a higher moment of the unknown function. One finds that the equation for a moment of any order involves moments of higher order. This procedure results in an infinite system of equations which must be solved simultaneously. To get a finite set of equations, unverified and often unjustified "closure" approximations are made. These closure approximations are the basis for what Keller (1964) calls "dishonest" methods for solving stochastic differential equations. The difficulty with the hierarchy

method is essentially that the average of the solution of a differential equation with stochastic coefficients is not necessarily the same as the solution of the averaged equation. Various closure approximations simply hide this difficulty. It has been shown by Adomian (1967, 1968) that hierarchy methods are valid only under special cases, such as small randomness. Even then it is often preferable to use approximate methods whose validity can be verified and wherein errors can be evaluated.

An important application of stochastic differential equation theory is the study of wave propagation in randomly turbulent media. Wave propagation in random media has been studied by Keller (1964), Hoffman (1964, 1959), Twersky (1964), Wheelon (1959), Bugnolo (1959, 1961), Lax (1951), Chernov (1960), Tatarski (1961), Mintzer (1953, 1954), Booker and Gordon (1950) and Booker (1959). More recent papers have been published in the special issue on partial coherence of IEEE Transactions on Antennas and Propagation (1967). Since the books by Chernov (1960) and Tatarski (1961) and the review article by Wheelon (1959) contain excellent reviews of the older literature and extensive bibliographies, only brief remarks are needed here. Nearly all studies to date of wave propagation in random media have assumed harmonic time dependence. In many cases, this "quasimonochromaticity" assumption is clearly not valid; in other cases, it needs verification. Examples of the first case are wave propagation through energetic media and interaction of electromagnetic waves in excited media. Also, small randomness is usually assumed from the outset, for example, in the derivation of the wave equation. It would be desirable to have a method where one applies restrictive assumptions

as late in the problem as feasible, so that it becomes clear exactly where the restrictive assumptions are needed and how they can be removed, if required for a particular problem. For these reasons, it is desirable to investigate wave propagation in the stochastic media from a more general point of view.

The rigorous mathematical background for the theory of stochastic equations has been reviewed in a survey paper by Bharucha-Reid (1964). He presents the basic definitions and theorems from probabilistic functional analysis that are used in the theory of random equations. The paper also reviews different classes of random equations such as random algebraic equations, random difference equations, random differential equations, and random integral equations. In a recent Ph. D. thesis, Anderson (1967) has studied in great detail Fredholm integral equations with stochastic forcing functions. Strand (1967) has studied existence and uniqueness of the ordinary stochastic differential equations. Goldstein (1967) has studied the sample function behavior of the second-order Itô processes. The operator theoretical treatment of this problem leads nonlinear semigroups of operators. Itô equations are very special stochastic differential equations and are not considered here.

The most general approach to stochastic differential equations or, in general, to stochastic operator equations, has been taken by Adomian (1961, 1963, 1964, 1967). He has developed the concept of a "stochastic Green's function". The integral kernel which expresses the desired statistical measures of the solution in terms of the corresponding statistical measure of the input and appropriate statistical measures of the stochastic coefficients is called the "stochastic

Green's function". The term "statistical measure" is used as a general term for the quantities that characterize stochastic processes. Expectations, spectral densities, and correlation functions are the statistical measures which are most widely used in physical problems. In the view of the fact that Green's functions have been widely used in applied mathematics, mathematical physics, linear system analysis, and electromagnetic theory, it can be expected that the concept of a stochastic Green's function may provide the unifying concept for a large number of diverse problems that are described by stochastic differential equations. The major problem in solving equations by means of stochastic Green's functions, as is the case with the ordinary Green's functions, is the problem of constructing the Green's function. The central problem of this dissertation is the construction and interpretation of stochastic Green's functions for various problems that arise in electrical engineering.

1.4 General Method of Solution and Specific Statement of the Problem

The ultimate goal of the solution of a stochastic differential equation is to express the statistical measures of the dependent variable in terms of the statistical measures of the forcing function and the stochastic coefficients. One approach to the problem is to determine all the orders of the multivariate probability distributions of the dependent variable. This gives the complete statistical description of the dependent variable and hence is the complete solution of the problem. This method is correct in principle, but is too difficult to be useful for most practical problems. In most applications, one is primarily interested in correlation functions,

spectral densities and mean values. These statistical measures can be computed from multivariate probability distributions but, in view of the complexity of the problem, it is desirable to seek expressions for the desired statistical measures of the dependent variable in terms of the same statistical measures of the forcing function and appropriate statistical measures of the stochastic coefficients. Another reason to avoid the intermediate steps is that, for the calculation of the multivariate probability distribution function, one needs to know all the multivariate probability distribution functions of the forcing function and stochastic coefficients. This information is not as frequently available as the second-order statistics (correlation functions, spectral densities, etc.). One needs to resort to more complicated experiments or unjustified assumptions to obtain the higher order information which is disregarded later anyway. Thus, a cleaner method is one which eliminates the intermediate steps of obtaining multivariate probability distributions.

Because of the difficulty of solving differential equations with stochastic coefficients, it is desirable to have a method which takes advantage of known solutions of the corresponding deterministic equation. To this end, one would seek an expression for the stochastic Green's function in terms of the known deterministic Green's function function. With this as motivation, the central problem of the dissertation can be rephrased in more precise terms as follows: Let $\mathcal{L}$ be an n^{th} order stochastic differential operator defined on some domain $t\in T$ and a probability space $(\Omega,\mathcal{F},P)$

$$\mathcal{F} = \sum_{\nu=0}^{n} a_\nu(t,\omega) \frac{d^\nu}{dt^\nu} , \tag{1.2}$$

where $a_\nu(t,\omega)$ are time varying random functions, $t\epsilon T$ and $\omega\epsilon\Omega$ on $(\Omega,\mathfrak{A},P)$. Let the operator $\mathcal{F}$ be separable into the sum of a deterministic operator L and a random operator R. In particular, let the random coefficients be of the form

$$a_\nu(t,\omega) = \beta_\nu(t) + \alpha_\nu(t,\omega) , \tag{1.3}$$

where $\beta_\nu(t)$ are deterministic functions of time and $\alpha_\nu(t,\omega)$ are stochastic processes { $\beta_\nu(t)$ can be either the ensemble average of $a_\nu(t,\omega)$ or some other convenient function of time }. For example, it may be possible to choose $\beta_\nu(t)$ so that the inversion of the deterministic differential operator L is simplified. It is assumed that L is an invertible differential operator; that is, the Green's function $G(t,\tau)$ for the differential operator L is known or can be constructed. One can associate with the operator $\mathcal{F}$ a stochastic differential equation

$$\mathcal{F}\, y\,(t,\omega) = x(t,\omega),\ t\epsilon T \text{ and } \omega\epsilon\Omega \text{ on } (\Omega,\mathfrak{A},P) . \tag{1.4}$$

The forcing function $x(t,\omega)$ can be either deterministic or random. For greater generality, let it be random. It is also assumed that $x(t,\omega)$ is statistically independent of the random coefficients. The problem is to find integral expressions for the statistical measures of $y(t,\omega)$ in terms of the same statistical measures of $x(t,\omega)$ and appropriate statistical measures of the stochastic coefficients. By the assumption that the Green's function for the deterministic operator L is known, the stochastic differential equation (1.4) is converted to a Volterra integral equation with a stochastic kernel and a stochastic forcing function. The integral equation can be solved by a Neumann series expansion and a resolvent kernel can be constructed. The stochastic Green's function can be expressed in terms of the resolvent kernel of the stochastic Volterra integral equation. This method of solution has the following advantages:

(i) Knowledge of the deterministic Green's function is used to construct the stochastic Green's function.

(ii) An iterative method of solution is used. The previously computed term is used to compute the next term and so on. Iteration can be stopped as soon as the remainder term reaches a prescribed value.

The main disadvantage of this method is that a large amount of labor may be required for the calculation of the resolvent kernel; but, on the other hand, note any simplifying assumptions can be made at places where their effects become clear. The solution of the stochastic integral equation is simplified if the kernel of the integral equation is degenerate. In many cases, the problem of

solving a stochastic differential equation, with a time invariant deterministic part, reduces to a problem of solving an integral equation with a degenerate kernel.

Chapter II contains general mathematical background, definitions and terminology which will be used in subsequent chapters. Some basic concepts from probability theory and from the theory of stochastic processes are presented. The last part of Chapter II contains a number of definitions and theorems from probabilistic functional analysis. These theorems give sufficient conditions for existence of a solution for the stochastic integral and differential equations. In reading this dissertation, one may go directly to Chapter III without loss of continuity and refer back to Chapter II for formal definitions and specific results as need arises.

The relations between resolvent kernels of Volterra integral equations and stochastic Green's functions are discussed in Chapter III. Both the Neumann series expansion and the degenerate kernel method are used for the construction of the resolvent kernel. In both cases, the computation is simplified if the state space formulation is used. The use of the state space formulation has the further advantage that it connects modern control system theory with this work.

In Chapter IV, both the degenerate kernel method and the Neumann series expansion are used to study propagation of a scalar wave function in a randomly time- and space-varying medium. The statistical measures of interest are the power spectral density and the coherence functions of the scalar wave function. The expressions for the spectral density will reveal the spectral spreading caused by a randomly time-varying medium. The usual quasimonochromatic

assumption, which is avoided in this thesis, fails to show the spreading of the power spectrum by the randomly time-varying medium.

Chapter V summarizes the results obtained in this thesis and discusses possible extensions to this work.

CHAPTER II

MATHEMATICAL BACKGROUND

2.1 Introduction

The purpose of this chapter is to establish the general mathematical background for the subsequent chapters, clarify terminology, give definitions, and to state useful theorems and inequalities. Since most material is readily available in standard books (Cramér and Leadbetter, 1967; Papoulis, 1965; Pugachev, 1965; Loève, 1963; Doob, 1953; Sveshnikov, 1966; and Blanc-Lapierre and Fortet, 1965) and review papers, (Moyal, 1949) the treatment is brief and necessarily incomplete. Most of the theorems are stated without proofs. First, some of the fundamental concepts from the probability theory and from the theory of stochastic processes are presented. Probability spaces, probability distributions, random variables and stochastic processes are defined. In the next section, various moments such as expected values (means), correlation functions, covariance functions and higher moments are discussed. A number of useful inequalities for the moments of stochastic processes are also presented. Then, in section 2.4, various concepts of stationarity are defined. The concepts of strict stationarity, wide-sense stationarity, reducibility to wide-sense stationarity are discussed. The following section (2.5) deals with the calculus of stochastic processes. The discussion begins with the definitions of various modes of convergence of the stochastic processes. Interrelations between these modes of convergence are briefly discussed. Using

the concept of convergence in the quadratic mean, quadratic mean continuity, quadratic mean differentiability and integrability are defined. The convergence of the iterative solutions of stochastic differential equations can be interpreted according to one of the modes of convergence of the stochastic processes. In section 2.6, spectral expansion of the stochastic processes is discussed and the power spectrum of wide-sense stationary stochastic processes is defined. In the following section the power spectrum of the non-stationary stochastic processes is defined in terms of a double Fourier transform. Section 2.8 develops some simple results which express the expected values and the covariance functions of stochastic processes which have been transformed by a linear deterministic operator. It is also shown that a Gaussian stochastic process remains Gaussian under linear transformation.

The final section of the chapter presents some definitions and theorems from probabilistic functional analysis. Probabilistic functional analysis provides a number of existence theorems for stochastic integral and differential equations.

We use the following convention. If a word is underlined in a sentence, that sentence serves to define the underlined word. The reader may go directly to Chapter III without loss of continuity and refer back to this chapter as need arises.

2.2 Probability Spaces, Random Variables and Stochastic Processes

Probability theory has its own terminology which is directly related to its intuitive background but, as a branch of mathematics, its concepts are expressible in terms of measure spaces and measurable

functions. To establish the basic concepts, let us consider a random experiment E with possible outcomes A,B,.... A,B,... are various observable events associated with the experiment E. A sure event, denoted by Ω, is an event which always occurs when the experiment E is performed. An impossible event, denoted by $\emptyset$, is an event which never occurs as an outcome of E. Both the sure event Ω and the impossible event $\emptyset$ are regarded as observable events. One sure "event" is the collection of all possible outcomes. For the processes we are considering that is the only sure event. Thus we use the symbol Ω to denote the whole space of events.

A space Ω with points ω, together with a σ-field $\mathfrak{X}$ of sets in Ω, and a probability measure P(A) defined on the sets A of $\mathfrak{X}$ constitutes a probability space denoted by $(\Omega,\mathfrak{X},P)$. A field $\mathfrak{X}$ of ω sets is called a Borel field, or a σ-field, if it includes all countable (finite or enumerable) unions and intersections of its sets. P(A) is said to define a probability distribution in Ω. P(A) is a function defined on all events or sets $A\in\mathfrak{X}$ and it has the following properties:

$$0 \leq P(A) \leq 1 \ , \tag{2.1a}$$

$$P(\emptyset) = 0 \ , \tag{2.1b}$$

$$P(\Omega) = 1 \ , \tag{2.1c}$$

and P(A) is countably additive; that is:

$$P\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} P(A_j) \ , \tag{2.1d}$$

for disjoint sets A_j ($A_j \cap A_k = \emptyset$ for $j \neq k$). The probability measure defined in this manner agrees with the usual intuitive concepts of the probability: the probability of an impossible event is zero, that of a sure event is one, and $r(A)$ is approximately equal the relative frequency of occurrence of the observable event A when the experiment E is performed a large number of times.

It is also assumed, for mathematical convenience, that the σ-field $\mathfrak{X}$ is completed with respect to the probability measure P(A). This means that all subsets of $\mathfrak{X}$ sets of P-measure zero are adjoined to $\mathfrak{X}$, and the smallest σ-field, including this extended family of sets, is formed. The completed σ-field has the property that if it includes a set A which has a P-measure zero, then it also includes every subset of A which will then also have a P-measure zero. The extension of P(A) to this completed σ-field is called a complete probability measure. This extension of probability measure defines probabilities for the events which may not be strictly observable, but this extension gives additional analytical freedom. Namely, a probability measure is defined for events that are obtained from the observable event by any set operations.

In the literature, the probability theory and the measure theory terms are used interchangeably. The correspondence between some of these terms is shown in the following table.

TABLE I

RELATIONSHIPS BETWEEN PROBABILITY THEORY
AND MEASURE THEORY

PROBABILITY THEORY	MEASURE THEORY
Probability space	Normed measure space
Sure event	Whole space Ω
Impossible event	Empty set
Event	Measurable set
Elementary event	Point ω belonging to the space
Probability	P-measure, normed measure
Almost sure, almost surely, a.s.	Almost everywhere P,a.e.P.
Random variable, r.v.	Measurable function
Expectation, statistical average mean, ensemble average, $\langle\ \rangle$	Integral $\int$

As stated in Table I, a random variable is a P-measurable function. Mathematically the definition of a random variable is simply the definition of measurability. A function $X(\omega)$ on Ω to $R = (-\infty, +\infty)$ is said to be measurable if, for every real number a, the set

$$\{\omega \in \Omega : X(\omega) \leq a\}$$

belongs to the σ-field $\mathfrak{X}$ (σ-field, measurable sets, events).

The measurable function $X(\omega)$ is called a random variable. Probability distributions are defined for all measurable sets or events. When needed, $X(\omega)$ may be allowed to become infinite or even undetermined on an ω set of P-measure zero. When the two random variables $x(\omega)$ and $y(\omega)$ are equal with probability one, written

$$P[x(\omega) = y(\omega)] = 1 \quad ,$$

they are called equivalent random variables. Equivalent random variables differ at most on an ω set of P-measure zero.

When $\zeta(\omega)$ is a random variable the probability $F(x) = P(\zeta \leq x)$ is a non-decreasing function of the real variable x. $F(x)$ is continuous to the right, and

$$\lim_{x \to -\infty} F(x) = 0 \qquad (2.2a)$$

$$\lim_{x \to \infty} F(x) = 1 \quad . \qquad (2.2b)$$

The function $F(x)$ is called the distribution function (or cumulative distribution function, d.f.) of the random variable ζ. The knowledge of $F(x)$ for all x determines the probability $P(\zeta \epsilon A)$ for every Borel set A. A is a subset of the real line R.

The random variables $\zeta_1, \zeta_2, \ldots \zeta_n$ will jointly induce a probability in n-dimensional Euclidean space R^n. The probability $F(x_1, x_2, \ldots x_n) = P(\zeta_1 \leq x_1, \zeta_2 \leq x_2, \ldots \zeta_n \leq x_n)$ is a nondecreasing function in each variable $x_1, x_2, \ldots x_n$, it is continuous on the right in each variable and

$$\lim_{x_i \to -\infty} F(x_1, x_2, \ldots x_i, \ldots x_n) = 0 \quad i = 1,2,\ldots n, \tag{2.3a}$$

and

$$\lim_{x_1, x_2, \ldots x_n \to \infty} F(x_1, x_2, \ldots x_n) = 1 \quad . \tag{2.3b}$$

$F(x_1, x_2, \ldots x_n)$ is the multivariate distribution function (joint distribution function, j.d.f.) of the random variables $\zeta_1, \zeta_2, \ldots \zeta_n$. Its knowledge determines the probability assigned to every Borel set of R^n by the n-dimensional distribution of the ζ_i. The random variables $\zeta_1, \zeta_2, \ldots \zeta_n$ can be thought of as components of an n-dimensional random vector

$$\bar{\zeta} = (\zeta_1, \zeta_2, \ldots \zeta_n) \quad .$$

The distribution function of the random vector $\bar{\zeta}$ is identical with the multivariate distribution of its components.

Now we turn to a generalization of the preceding concepts of random variables and their distribution functions. That is, we shall consider an arbitrary family $\{X(t,\omega)\}$ of random variables where t runs through some index set T. If T consists of a single point we have a single random variable; a finite T-set corresponds to the finite family of random variables. When T is an interval of real numbers, the family $\{X(t,\omega):t\epsilon T\}$ is called a continuous-parameter stochastic process or a random function. Physical systems subject to random influences can be described by stochastic processes. Another way of looking at stochastic processes is the following. A random variable X is a set function on the sample space Ω; that is, a random variable X assigns a number $X(\omega)$ to every $\omega\epsilon\Omega$. In the case of a stochastic process, for every $\omega\epsilon\Omega$, a function $X(t,\omega)$ is assigned. Hence, we are dealing with an ensemble or a family of functions. $X(t,\omega)$ is often a function of time, but it may very well be a function of any other quantity such as position. With the preceding discussion as motivation, a rigorous definition of a stochastic process is now given. Let a probability space $(\Omega,\mathfrak{T},P)$ and a parameter set T be given. A stochastic process is a finite real valued function $X(t,\omega)$ which, for every fixed $t\epsilon T$, is a measurable function of $\omega\epsilon\Omega$. A stochastic vector process is an n-dimensional vector whose components are stochastic process; that is:

$$\bar{X}(t,\omega) = \{X_1(t,\omega), X_2(t,\omega) \cdots X_n(t,\omega)\}$$

A particular vector process, with n = 2, is a complex-valued stochastic process $X(t,\omega) + i\,Y(t,\omega)$. There are several ways to interpret stochastic processes. Stochastic processes can be considered as a collection, or an ensemble or a family of functions. Each particular function in the collection is called a sample function, a realization of a stochastic process, or a representative function. If t is fixed at any particular t_o, then $X(t_o,\omega)$ is simply a random variable. On the other hand, if ω is fixed, then $X(t,\omega_o)$ is simply a sample function of t. If both ω and t are fixed we have a single number. The stochastic process can be very complicated or quite simple. An example of a complicated process is Brownian motion. The sample functions are the coordinates of a particular particle which is in irregular motion in a liquid or gas. The irregular motion of the particle is caused by the random impacts with the particles in the surrounding medium. If the particle has negligible bonds with the medium except at times of impact, almost all Brownian movement sample functions are of unbounded variation! (Doob, 1953; p 395). An example of a simple process is the output of a signal generator

$$X(t,\omega) = a(\omega)\cos\left(2\pi\lambda(\omega)t + \theta(\omega)\right) \quad , \qquad (2.4)$$

where $a(\omega)$ is the random amplitude, $\lambda(\omega)$ is the random frequency and $\theta(\omega)$ is the random phase. This is precisely what is expected in a usual experimental situation: the amplitude, frequency and phase of the signal generator are only known within the limits of the measurement accuracy. $X(t,\omega)$ as given by (2.4) is a stochastic process but its sample functions have very simple analytic properties. From the mathematical point of view, a stochastic process is a function

$X(t,\omega)$ of two variables t and ω. The domains of definition of t and ω are the sets T and Ω.

For t fixed at t_1, we have a random variable $X(t_1,\omega)$ with a first-order distribution function for the stochastic $X(t,\omega)$ given by

$$F(x;t_1) = P\{X(t_1,\omega) \leq x\} \quad . \tag{2.5}$$

For an arbitrary finite set of values of t, t_1, t_2 --- t_n, we have the corresponding random variables $X(t_1,\omega)$, $X(t_2,\omega)$ --- $X(t_n,\omega)$ with the n-dimensional joint distribution function

$$F(x_1,x_2,\cdots,x_n;t_1,t_2,\cdots,t_n) = P\{X(t_1,\omega) \leq x_1,\cdots,X(t_n,\omega) \leq x_n\} \quad . \tag{2.6}$$

Clearly a stochastic process is completely specified if the distribution functions are known for all n . In many physical problems, such a complete knowledge is not available and one must be satisfied with a knowledge of say its second-order statistics (statistics calculated from the second-order distribution functions, for example, correlations, power spectrum). Second order properties are discussed in the next section.

If the distribution functions are differentiable, the probability density functions exist and are given, for the first-order distribution function by

$$f(x,t) = \frac{\partial F(x,t)}{\partial x} \tag{2.7a}$$

and for the n-dimensional joint distribution function, by

$$f(x_1,x_2,\cdots,x_n;t_1,t_2,\cdots t_n) = \frac{\partial^n F(x_1,x_2,\cdots x_n;t_1,t_2,\cdots t_n)}{\partial x_1\,\partial x_2 \cdots \partial x_n} \tag{2.7b}$$

2.3 Moments of Stochastic Processes

As was stated before, a stochastic process is completely specified if its distribution functions are known for all finite n . In many physical problems, such complete knowledge is not available or it is excessively complicated or costly to measure these distribution functions experimentally. An alternative specification is to use various moments (defined below) of the stochastic process. The first two moments have found wide use in the communication and control system theory. The expected value (ensemble average, mathematical expectation, statistical average), of any function g(x,t) of a stochastic process is given by

$$\langle g(x(\omega),t) \rangle = \int_{-\infty}^{\infty} g(x,t)\, d\, F(x,t) \,. \tag{2.8}$$

In this dissertation, the symbol $\langle g(x,t) \rangle$ is used to denote the ensemble averaging. Equation (2.8) is a Riemann-Stieltjes integral. If the probability density function, f(x,t) of $x(t,\omega)$ exists, the expected value is given by

$$\langle g(x(\omega),t) \rangle = \int_{-\infty}^{\infty} g(x,t)\, f(x,t)\, dx \,, \tag{2.9}$$

where (2.9) is the usual Riemann integral. The particular expected value of x, $< x(t,\omega) > = \mu_x(t)$, is the mean or first moment, and $< (x(t,\omega) - \mu(t))^2 > = \sigma^2(t)$ the variance of the stochastic process. The function $< x(t,\omega)^n >$ is the n-th moment of $x(t,\omega)$ and $< (x(t,\omega) - \mu(t))^n >$ is the n-th central moment of $x(t,\omega)$.

If $X(t,\omega)$ and $Y(t,\omega)$ are independent stochastic processes, then, by definition of independence,

$$F(x,y;t_1,t_2) = F(x;t_1)\, F(y;t_2) \tag{2.10}$$

and

$$< X(t_1;\omega)\, Y(t_2;\omega) > = \int_{-\infty}^{\infty}\int_{\infty}^{\infty} g(x;t_1)\, h(y;t_2)\, dF(x,y;t_1,t_2)$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x;t_1)\, h(y;t_2)\, dF(x;t_1)\, dF(y;t_2)$$

$$= \int_{-\infty}^{\infty} g(x;t_1)\, dF(x;t_1) \int_{-\infty}^{\infty} h(y;t_2)\, dF(y;t_2)$$

$$= < X(t_1,\omega) > < Y(t_2,\omega) > \quad . \tag{2.11}$$

The above property of the expectation operator will be used frequently to separate the ensemble averages of statistically independent stochastic processes. If either $< X(t_1,\omega) >$ or $< Y(t_2;\omega) >$ is zero and $X(t_1,\omega)$ and $Y(t_2,\omega)$ are statistically independent, then obviously

$$< X(t_1,\omega)\ Y(t_2,\omega) > = < X(t_1,\omega) > < Y(t_2,\omega) > = 0 \quad . \tag{2.12}$$

The stochastic processes $X(t,\omega)$ and $Y(t,\omega)$ are said to be uncorrelated if

$$< X(t,\omega)\ Y(t,\omega) > = < X(t,\omega) > < Y(t,\omega) > \quad .$$

Statistical independence implies that random processes are uncorrelated, but not vice versa. If

$$< X\ (t,\omega)\ Y(t,\omega) > = 0 \quad , \tag{2.13}$$

the stochastic processes are said to be orthogonal. Orthogonality is used frequently to simplify expressions involving stochastic processes.

The most widely used moments are the first and second moments, the mean and correlation functions, respectively. The theory based on the first two moments is called the second-order theory or the correlation theory. This dissertation deals principally with the second-order theory of stochastic differential equations. Unless otherwise stated, it is assumed that second moments of the stochastic processes exist.

For a complex process, the autocorrelation function is defined by

$$R_{XX}(t_1,t_2) = < X(t_1,\omega)\ X^*(t_2,\omega) > , \tag{2.14}$$

where * denotes the complex conjugate. The autocovariance is defined by

$$C_{XX}(t_1,t_2) = < [X(t_1,\omega) - \mu(t_1)]\ [X(t_2,\omega) - \mu(t_2)]^* >$$

$$= R(t_1,t_2) - \mu(t_1)\ \mu^*(t_2) \quad ,$$

where $\mu(t_1) = \langle X(t_1,\omega) \rangle$ and

$$\mu^*(t_2) = \langle X^*(t_2,\omega) \rangle \quad . \tag{2.15}$$

When the means $\mu(t_1) = \mu(t_2) = 0$, the autocorrelation and the autocovariance are equal. Often, it is convenient to work with zero mean processes for this reason. In such cases, the terms correlation and covariance can be used interchangeably.

The cross-correlation of two stochastic processes is defined by

$$R_{XY}(t_1,t_2) = \langle X(t_1,\omega)\, Y^*(t_2,\omega) \rangle \ , \tag{2.16}$$

and their cross-covariance by

$$C_{XY}(t_1,t_2) = R_{XY}(t_1,t_2) - \mu_X(t_1)\, \mu_Y(t_2) \quad . \tag{2.17}$$

A number of useful inequalities can be derived by considering:

$$[X(t_1,\omega) + k\, Y(t_2,\omega)]\, [X(t_1,\omega) + k\, Y(t_2,\omega)]^*$$

$$= R_{XX}(t_1,t_1) + k\, [R_{XY}(t_1,t_2) + R_{YX}(t_2,t_1)] + k^2\, R_{YY}(t_2,t_2) \geq 0 \quad . \tag{2.18}$$

The above quadratid in k is nonnegative for every k; hence k must have no real roots. This means that the discriminant of the quadratic in k must be nonpositive. Therefore, we have the following inequality;

$$[R_{XY}(t_1,t_2) + R_{YX}(t_2,t_1)]^2 \leq 4\, R_{XX}(t_1,t_1)\, R_{YY}(t_2,t_2) \ . \tag{2.19}$$

If $X(t,\omega)$ and $Y(t,\omega)$ are real processes, (2.19) simplifies to:

$$R_{XY}^{\ 2}(t_1,t_2) \leq R_{XX}(t_1,t_1)\, R_{YY}(t_2,t_2) \tag{2.20}$$

If $X(t,\omega) = Y(t,\omega)$ then we have a special case of (2.20)

$$R_{XX}^{\ 2}(t_1,t_2) \leq R_{XX}(t_1,t_1)\, R_{XX}(t_2,t_2) \ . \tag{2.21}$$

The normalized correlation coefficient is defined by

$$\rho_{XX}(t_1,t_2) = \frac{R_{XX}(t_1,t_2)}{[R_{XX}(t_1,t_1)\, R_{XX}(t_2,t_2)]^{\frac{1}{2}}} \ , \tag{2.22}$$

and cross-correlation coefficient by

$$\rho_{XY} = \frac{R_{XY}(t_1,t_2)}{[R_{XX}(t_1,t_1)\, R_{YY}(t_2,t_2)]^{\frac{1}{2}}} \ . \tag{2.23}$$

From (2.22) and (2.20) we have

$$0 \leq |\rho_{XX}| \leq 1 \tag{2.24}$$

$$0 \leq |\rho_{XY}| \leq 1 \tag{2.25}$$

The second order statistics of an n dimensional stochastic vector are given by a covariance (or correlation) matrix. The <u>covariance matrix</u> of an n dimensional stochastic vector $\bar{X}(t,\omega) = [X_1(t,\omega), X_2(t,\omega) \text{ --- } X_n(t,\omega)]$ is a n x n matrix with the elements $C_{x_i x_j}(t_1,t_2)$. If all the elements of a stochastic vector have zero mean, the covariance matrix can be obtained by computing $< \bar{X}(t_1;\omega)\, \bar{X}^{\dagger}(t_2,\omega) >$, where $\bar{X}(t_1;\omega)$ is an n dimensional column vector and † denotes the complex conjugate transpose (hermetian conjugate). In section 3.6, the solution of a control system problem with a stochastic state transition matrix is expressed in terms of the covariance matrix of the state variables. The terms "state transition matrix" and "state variables" are defined in section 3.6.

Another term used for the correlation or covariance functions (both terms may be used interchangeably since we are talking about zero mean stochastic processes) in partial coherence theory is the mutual coherence function. The <u>mutual coherence function</u> is simply the following ensemble average:

$$C_{Y_1 Y_2}(t_1,t_2,;P_1,P_2) = < Y_1(t_1,P_1,\omega)\, Y_2^{*}(t_2,P_2,\omega) > , \tag{2.26}$$

where $Y_1(t_1,P_1,\omega)$ and $Y_2(t_2,P_2,\omega)$ are complex field disturbances at two points P_1 and P_2. The mutual coherence function defined in this manner is the ensemble coherence function. If time averaging is used, one has a time coherence function. Both the ensemble coherence functions and time coherence functions have been used in statistical optics and in general stochastic electromagnetic theory. The ensemble coherence functions are discussed further in Chapter IV.

A number of basic inequalities and properties can be established for stochastic processes. Some of the more useful ones are stated without proofs. Proofs and more complete discussion is found in Loéve, (1963, page 156).

1. If for some $m > 0$, $< |X(t,\omega)|^m > < \infty$, then $< |X(t,\omega)|^n >$ is finite for $n \leq m$, and $< X^k(t,\omega) >$ exists and is finite for $k \leq m$.

2. Hölder inequality

   For any two processes X and Y or the finite moments,
   $< |XY| > \leq < |X|^{\frac{1}{m}} >^m < |Y|^{\frac{1}{n}} >^n$, where $m > 1$ and $\frac{1}{m} + \frac{1}{n} = 1$.

   (2.27)

3. A special case of the Hölder inequality is the Schwarz inequality $< |XY| >^2 \leq < |X|^2 > < |Y|^2 >$.

   (2.28)

4. Minkowski inequality

   Under the same hypothesis as 2,
   $< |X + Y|^n >^{\frac{1}{n}} \leq < |X|^n >^{\frac{1}{n}} + < |Y|^n >^{\frac{1}{n}}$

   (2.29)

5. If $< |X|^n >$ exists for each n, then $< |X|^n >^{\frac{1}{n}}$ is a nondecreasing function of n.

2.4 Stationarity Concepts

In some applications, the statistical properties of stochastic processes exhibit some invariance under translation of time, space or some other independent variable. As more or less stringent conditions of invariance of the statistics are imposed, different types of stationarity may be defined. A stochastic process $X(t,\omega)$ is said to be strictly stationary (stationary, stationary in the strict sense, statistically homogeneous) if the whole family of its finite-dimensional distributions are invariant under a translation in the parameter t; i.e.,

$$F(x_1,x_2,\ldots x_n;\ t_1,t_2,\ldots t_n) = F(x_1,x_2,\ldots x_n;\ t_1 + h,\ t_2 + h,\ldots t_n + h) \tag{2.30}$$

for any n, $t_1,t_2,\ldots,t_n$, and h. The statistics of the strictly stationary stochastic processes are not affected by the choice of time origin. The processes $X(t,\omega)$ and $Y(t,\omega)$ are jointly stationary (in the strict sense) if the joint distributions are invariant under a translation in the parameter t. A complex process

$$Z(t,\omega) = X(t,\omega) + i\,Y(t,\omega) \tag{2.31}$$

is said to be stationary if $X(t,\omega)$ and $Y(t,\omega)$ are jointly stationary. A process $X(t,\omega)$ is stationary of order k if (2.30) is true only for $n \leq k$. A process which is stationary of order 2 is also a wide-sense stationary (weakly stationary) process. A stochastic process is a wide-sense stationary process if

$$\langle X(t,\omega) \rangle = \mu = \text{constant},$$

and

$$< X(t_1;\omega)\, X^*(t_2;\omega) > = < X(t_1;\omega)\, X^*(t_1 + \tau;\, \omega) > = R_{XX}(\tau) \quad . \tag{2.32}$$

The wide-sense stationarity involves only the first two moments. For this reason, in a second-order theory only the weaker assumption of wide-sense stationarity is used instead of the stronger assumption of strict stationarity. Two processes are <u>jointly stationary</u> in the <u>wide sense</u> if each is stationary in the wide sense and their cross-correlation depends only on $t_2 - t_1$:

$$R_{XY}(\tau) = < X(t_1;\omega)\, Y^*(t_1 + \tau;\, \omega) > \quad . \tag{2.33}$$

One often meets stochastic processes which can be expressed comparatively simply in terms of wide-sense stationary stochastic processes. Such stochastic processes are said to be <u>reducible</u> to wide-sense stationary processes. An example of such a process is

$$Y(t,\omega) = f(t)\, X(t,\omega) + g(t), \tag{2.34}$$

where $X(t,\omega)$ is a wide-sense stationary stochastic process, $f(t)$ and $g(t)$ are real functions. The mean of $Y(t,\omega)$ is

$$\mu_Y(t) = f(t)\, \mu_X + g(t), \tag{2.35}$$

the covariance of $Y(t,\omega)$ is

$$C_{YY}(t_1,t_2) =$$

$$= \left\langle \left(f(t_1)X(t_1,\omega)+g(t_1)-f(t_1)\mu_X-g(t_1) \right) \left(f(t_2)X(t_2,\omega)+g(t_2)-f(t_2)\mu_X-g(t_2) \right)^* \right\rangle$$

$$= f(t_1)\, f(t_2)\, R_{XX}(t_2-t_1) - 2f(t_1)\, |\mu_X|^2 =$$

$$= f(t_1)\, f(t_2)\, C_{XX}(t_2-t_1), \tag{2.36}$$

and the variance of $Y(t,\omega)$ is

$$C_{YY}(t,t) = f^2(t)\, C_{XX}(0) \quad . \tag{2.37}$$

The normalized covariance function of $Y(t,\omega)$ is

$$c_{YY}(t_1,t_2) = \frac{f(t_1)\, f(t_2)\, C_{XX}(t_2-t_1)}{\left[f^2(t_1)\, f^2(t_2)\, C_{XX}^{\ 2}(0) \right]^{\frac{1}{2}}} = c_{XX}(t_2-t_1), \tag{2.38}$$

which is the same as the normalized covariance function of the wide-sense stationary process $X(t,\omega)$.

The important concepts of time averages, time autocorrelation functions and egodicity are not discussed here because, in this dissertation, ensemble statistics are used exclusively.

2.5 Calculus of Stochastic Processes

In order to study the analytical properties of stochastic processes, the concept of convergence must be defined first. Suppose a sequence of random variables $x_1(\omega)$, $x_2(\omega)$, ..., is given and all

the random variables $x_n(\omega)$ are defined on the same probability space Ω. Let $x(\omega)$ be another random variable defined on the same probability space. One could define convergence in the same manner as it is defined in deterministic analysis, requiring that the sequence $x_n(\omega)$ converges to $x(\omega)$ for every realization of the sequence. Such a definition of convergence is too restrictive. In stochastic theory, it is convenient to allow less stringent modes of convergence. Three most important modes of convergence of $x_n(\omega)$ to the limit $s(\omega)$ as $n \to \infty$ are:*

1) $x_n(\omega)$ converges to $x(\omega)$ almost everywhere P-measure (a.e.P), or with probability one, or almost sure (a.s.), if

$$P(x_n(\omega) \to x(\omega)) = 1.$$

2) $x_n(\omega)$ converges to $x(\omega)$ in quadratic mean (q.m.), or in mean square (m.s.) or limit in the mean (l.i.m.), if

$$\langle |x_n(\omega) - x(\omega)|^2 \rangle \to 0$$

3) $x_n(\omega)$ converges to $x(\omega)$ in probability, or in P-measure, if, for every $\epsilon > 0$, $P(|x_n(\omega) - x(\omega)| > \epsilon) \to 0$.

In addition to these three, a fourth one, is sometimes used.

4) $x_n(\omega)$ converges to $x(\omega)$ in distribution function, if at every point of continuity of $F(x)$, $F(x_n) \to F(x)$.

The first three modes of convergence are analogous to the corresponding modes of convergence in measure theory. One can show that almost everywhere (a.e.P.) and quadratic mean (q.m.) convergence imply

---

*Cramér and Leadbetter, 1967

convergence in probability and convergence in probability implies convergence in distribution function (Papoulis, 1965). The relations between these four modes of convergence are shown in Figure 2.1. In this dissertation, almost everywhere and quadratic mean convergence are used most often. Using the definition of quadratic mean convergence, quadratic mean continuity and quadratic mean differentiation can be defined. A second order stochastic process is <u>continuous in quadratic mean</u> at $t\epsilon T$ if $X(t + h;\omega)$ converges to $X(t,\omega)$ in q.m. as $h \to 0$, $t + h\epsilon T$. A second-order stochastic process $X(t,\omega)$ on T has a <u>derivative in quadratic mean</u> $\frac{dX(t,\omega)}{dt}$, at $t\epsilon T$ if

$$\frac{X(t + h, \omega) - X(t,\omega)}{h} \xrightarrow{\text{q.m.}} \frac{dX(t,\omega)}{dt}, \text{ as } h \to 0, (t + h)\epsilon T. \tag{2.39}$$

By using almost everywhere convergence instead of quadratic mean convergence, almost everywhere continuity and differentiability can be defined. Higher order derivatives and partial derivatives in quadratic mean can be defined similarly to (2.39). The quadratic mean Riemann, Riemann Stieltjes and Lebesgue integrals can be defined by using the quadratic mean convergence of the approximating sums which are used in the definitions of these integrals. The ordinary formal properties, such as additivity and integration by parts, hold for integrals in quadratic mean under the appropriate conditions (Loève, 1963). The following theorem, which is stated and proved by Loève (1963), will be later used to find expectations of iterated integrals.

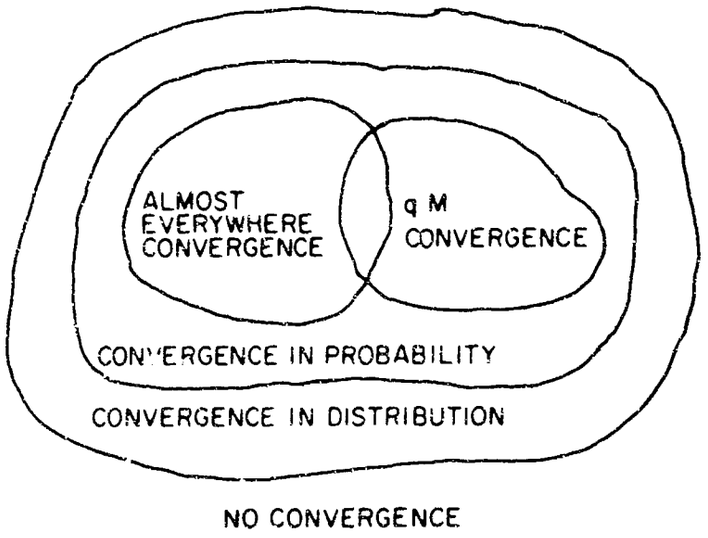


Figure 2.1 - Comparison of Various Modes of Convergence

Theorem 1. Let the second order stochastic processes $X(t,\omega)$ [with covariance function $C_{XX}(t,t')$] be independent of the second-order increment function $\Delta Y(t',\omega)$ (with covariance function $\Delta\Delta' C_{YY}(t,t')$ on an interval I X I, where I = [a,b] is a finite or infinite interval). Then,

$$\int_I X(t,\omega)\, dY(t,\omega) \text{ exists if, and only if,}$$

$$\int_I \int_I C_{XX}(t,t')\, dd'\, C_{YY}(t,t') \text{ exists; also, if the integrals}$$

in quadratic mean which appear below exist, then

$$< \int_I X(t,\omega)\, dY(t,\omega) \int_{I'} X^*(t',\omega)\, dY^*(t',\omega) >$$

$$= \int_I \int_{I'} < X(t,\omega)\, X^*(t',\omega) >\, dd'\, C_{yy}(t,t') \quad . \tag{2.40}$$

The double integrals are the usual Riemann-Stieltjes integrals. The independence condition of this theorem is fulfilled when the stochastic process $X(t,\omega)$ and $Y(t,\omega)$ are independent or when either $X(t,\omega)$ or $\Delta Y(t,\omega)$ degenerate into deterministic functions. The independence condition can be suppressed altogether when the elements of the double integrals are replaced by $dd < X(t,\omega)\, X^*(t',\omega)\, Y(t,\omega)\, Y^*(t',\omega) >$.

The convergence of the iterative solutions of stochastic differential equations can be interpreted according to one of the modes of convergence. Probabilistic functional analysis, as discussed in section 2.9, uses the almost everywhere P convergence (almost sure convergence). The convergence in quadratic mean (mean square convergence) also has found wide use in applications such as communication and control system theory.

2.6 Expansions of Stochastic Processes and Spectral Theory

Analytical operations with stochastic processes can be simplified if they can be represented as linear combinations of orthogonal random variables. A stochastic process $x(t,\omega)$ can frequently be expressed either as an infinite series

$$X(t,\omega) = \mu_x(t) + \sum_{k=1}^{\infty} X_k(\omega)\, g_k(t) \quad , \tag{2.41}$$

or as an integral

$$x(t,\omega) = \mu_x(t) + \int_{-\infty}^{\infty} X(s,\omega)\, g(t,s)\, ds \quad . \tag{2.42}$$

In (2.41) $X_k(\omega)$ are orthogonal random variables and $g_k(t)$ are deterministic functions. In (2.42) $X(s,\omega)$ is a random function of the parameter s, $g(t,s)$ is a deterministic function of time and s. The random variables $X_k(\omega)$ satisfy the following orthogonality condition:

$$\begin{aligned} \langle X_i(\omega)\, X_j(\omega) \rangle &= 0 \quad \text{for } i \neq j \\ &= g_i^{\,2} \quad \text{for } i = j. \end{aligned} \tag{2.43}$$

The corresponding orthogonality condition for the integral expansion (2.42), written as the following Riemann-Stieltjes integral,

$$x(t,\omega) = \mu_x(t) + \int_{-\infty}^{\infty} g(t,s)\, d_s\, X(s,\omega), \tag{2.44a}$$

is

$$< d_s\, X(s,\omega) > = 0 \tag{2.44b}$$

and

$$< d_{s_1}\, X(s_1;\omega)\, d_{s_2}\, X^*(s_2;\omega) > = \Phi_{xx}(s_1)\, \delta(s_1 - s_2)\, ds_1\, ds_2 \,. \tag{2.44c}$$

The expansions (2.41) and (2.42) have been called in the literature by a number of different names. They are known as the canonical expansions (Pugachev, 1965), orthogonal decompositions (Loève, 1963) or as the Karhunen-Loève expansion. In Chapter IV, the integral expansion (2.42) is used to study a wave propagation problem in a randomly time- and space-varying medium. The dielectric permittivity is assumed to be a wide-sense stationary stochastic process. For this reason, the following discussion is restricted to the integral expansion. For simplicity, it is assumed that $\mu_x(t) = 0$. For the wide-sense stationary zero mean process, the integral expansion (2.42) is (Pugachev, 1965, p. 239):

$$x(t,\omega) = \int_{-\infty}^{\infty} X(s,\omega)\, e^{ist}\, ds \,, \tag{2.45a}$$

where $X(s,\omega)$ is given by

$$X(s,\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} x(t,\omega)\, e^{-ist}\, dt \,. \tag{2.45b}$$

The integral expansion (2.45a) is also known as the spectral representation of the wide-sense stationary stochastic processes, and (2.45a) should be written as a Riemann-Stieltjes integral

$$x(t,\omega) = \int_{-\infty}^{\infty} e^{ist}\, d_s\, X(s,\omega); \tag{2.46}$$

however it is more convenient to operate with the integral expansion as it is expressed by (2.45a). In such case, it must be kept in mind that $X(s,\omega)$ may be a generalized function. If $X(s,\omega)$ is given by (2.45b), then we have, for the wide-sense stationary processes,

$$\begin{aligned} &< X(s_1,\omega)\, X^*(s_2,\omega) > = \\ &= \left(\frac{1}{2\pi}\right)^2 \int_{-\infty}^{\infty}\int < x(t_1;\omega)\, x^*(t_2;\omega) > e^{-is_1t_1}\, e^{is_2t_2}\, dt_1\, dt_2 \\ &= \left(\frac{1}{2\pi}\right) \int_{-\infty}^{\infty}\int R_{xx}(t_2 - t_1)\, e^{-is_1t_1}\, e^{is_2t_2}\, dt_1\, dt_2 \,. \end{aligned} \tag{2.47}$$

Making the change of variable $t_2 - t_1 = \tau$, (2.47) becomes

$$\begin{aligned} &< X(s_1,\omega)\, X^*(s_2,\omega) > = \\ &= \left(\frac{1}{2\pi}\right)^2 \int_{-\infty}^{\infty}\int R_{xx}(\tau)\, e^{-s_1t_1}\, e^{is_2\tau}\, d\tau\, dt_1 \,. \end{aligned} \tag{2.48}$$

But,

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{it_1(s_2 - s_1)}\, dt_1 = \delta(s_2 - s_1) \quad . \tag{2.49}$$

Hence

$$\begin{aligned} \langle X(s_1;\omega)\, X^*(s_2;\omega)\rangle &= \delta(s_2 - s_1)\ \frac{1}{2\pi}\int_{-\infty}^{\infty} R_{xx}(\tau)\, e^{is_2\tau}\, d\tau \\ &= \delta(s_2 - s_1)\ \frac{1}{2\pi}\int_{-\infty}^{\infty} R_{xx}(\tau)\, e^{-is_1\tau}\, d\tau \ . \end{aligned} \tag{2.50}$$

This relation (2.50) will be used often in Chapter IV. The <u>power spectral density</u> $\Phi_{xx}(s)$ (power spectrum) of a wide-sense stationary stochastic process is, by definition,

$$\Phi_{xx}(s) = \frac{1}{2\pi}\int_{-\infty}^{\infty} R_{xx}(\tau)\, e^{-is\tau}\, d\tau \quad . \tag{2.51}$$

The spectral density and the correlation functions are Fourier transform pairs of one another, i.e.;

$$R_{xx}(\tau) = \int_{-\infty}^{\infty} \Phi_{xx}(s)\ e^{is\tau}\, ds \quad . \tag{2.52}$$

There is no universal agreement whether the factor $1/2\pi$ is placed in front of (2.51) or in front of (2.52). The variance of the stochastic process $X(t,\omega)$ can be found from the power spectral density by:

$$R_{xx}(0) = \sigma^2 = \int_{-\infty}^{\infty} \Phi_{xx}(s)\ ds \quad , \tag{2.53}$$

From the assumption that $X(t,\omega)$ is a second order process, that is σ^2 is finite, we have:

$$\int_{-\infty}^{\infty} \Phi_{xx}(s)\, ds < \infty \quad , \tag{2.54}$$

For real stochastic processes, $R_{xx}(\tau)$ is an even function of τ, and it can be easily shown by making a change of variable $\tau = -\tau_1$ in (2.51) that

$$\Phi_{xx}(s) = \Phi_{xx}(-s) \quad . \tag{2.55}$$

The <u>spectral function</u> is defined by:

$$S_{xx}(s_1) = \int_{-\infty}^{s_1} \Phi_{xx}(s)\, ds \quad . \tag{2.56}$$

It can be shown that $S_{xx}(s_1)$ is a real nondecreasing bounded function of its argument s_1 (Khintchine, 1934). From this property it follows that

$$\Phi_{xx}(s) \geq 0 \quad . \tag{2.57}$$

The <u>mutual spectral density</u> of jointly wide-sense stationary second order stochastic processes $X(t,\omega)$ and $Y(t,\omega)$ is given by

$$\Phi_{xy}(s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} R_{xy}(\tau)\, e^{-is\tau}\, d\tau \quad . \tag{2.58}$$

In this section the discussion of the power spectrum was limited to the wide-sense stationary processes. In the next section, the definition of the power spectral density is generalized to be applicable to the nonstationary processes. This generalization is needed for the discussion of nonstationary solutions of the scalar wave equation.

2.7 Spectrum of Nonstationary Processes*

When the stochastic process $X(t,\omega)$ is not wide-sense stationary, its autocorrelation function is $R_{XX}(t_1,t_2)$; that is, the autocorrelation function is a function of t_1 and t_2 rather than just $t_2 - t_1$. For this reason the power spectral density for a nonstationary process is given by the double Fourier transform of $R_{XX}(t_1,t_2)$:

$$\Phi_{XX}(s_1,s_2) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty}\int R_{XX}(t_1,t_2)\, e^{+is_1t_1 - is_2t_2}\, dt_1\, dt_2, \tag{2.59}$$

with the inversion formula

$$R(t_1,t_2) = \int_{-\infty}^{\infty}\int \Phi_{XX}(s_1,s_2)\, e^{-is_1t_1 + is_2t_2}\, ds_1\, ds_2 . \tag{2.60}$$

From the fact that $R_{XX}(t_1,t_2) = R_{XX}^*(t_2,t_1)$ and from (2.59),

$$\Phi_{XX}(s_1,s_2) = \Phi_{XX}^*(s_2,s_1) . \tag{2.61}$$

---

*Papoulis (1965), and Blanc-Lapierre and Fortet (1965).

If $R(t_1,t_2)$ is real, then

$$\Phi_{XX}(-s_1,-s_2) = \Phi_{XX}^{*}(s_1,s_2) \quad . \tag{2.62}$$

It is easy to see that the general double transform (2.59) reduces to our previous result of $X(t,\omega)$ wide-sense stationary, i.e.:

$$\Phi_{XX}(s_1,s_2) = \delta(s_2-s_1) \quad \frac{1}{2\pi} \int_{-\infty}^{\infty} R_{XX}(\tau)\, e^{-is_1\tau}\, d\tau$$

$$= \delta(s_2-s_1) \quad \Phi_{XX}(s_1) \quad . \tag{2.63}$$

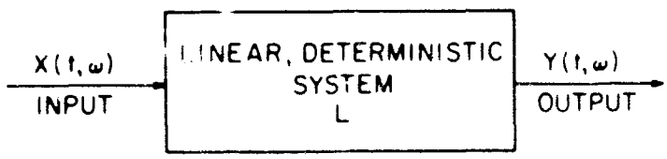
On the other hand, if $\Phi_{XX}(s_1,s_2)$ is equal to $\delta(s_2-s_1)\ \Phi_{XX}(s_1)$, the inversion formula (2.60) becomes

$$R_{XX}(t_1,t_2) = \int_{-\infty}^{\infty}\int \Phi_{XX}(s_1) \quad \delta(s_2-s_1)\, e^{-is_1t_1 + is_2t_2}\, ds_1\, ds_2$$

$$= \int_{-\infty}^{\infty}\int \Phi_{XX}(s_1)\, e^{is_1(t_2-t_1)}\, ds_1$$

$$= R_{XX}(t_2-t_1) = R_{XX}(\tau), \tag{2.64}$$

which shows that $X(t,\omega)$ is a wide-sense stationary stochastic process. The above results can be summarized as the following theorem.



$$Y(t,\omega) = LX(t,\omega)$$

Figure 2.2 - Linear Deterministic Transformation of a Stochastic Process

function of the output process, if it can be assumed that the expected value operator and the linear operator can be interchanged. The expected value operator and the linear operator are interchangeable for practically almost all deterministic linear operations. For example, it can be shown that the differential and integral operators are interchangeable with the expected value operations, (Pugachev, 1965, p. 387). Thus, the expected value of $Y(t,\omega)$ is

$$< Y(t,\omega) > = < L X(\tau,\omega) > = L < X(\tau,\omega) >, \qquad (2.67a)$$

or

$$\mu_Y(t) = L\mu_X(\tau). \qquad (2.67b)$$

If (2.67) is subtracted from (2.65), we have, according to (2.66)

$$Y_o(t,\omega) = L X_o(\tau,\omega) , \qquad (2.68)$$

where the subscript o denotes that $Y_o(t,\omega)$ and $X_o(\tau,\omega)$ are processes with zero means. For zero mean processes, both the covariance function and correlation functions of $Y_o(t,\omega)$ are given by the same expressions

$$R_{Y_o Y_o}(t_1,t_2) = \langle Y_o(t_1,\omega)\, Y_o^*(t_2,\omega) \rangle$$

$$= \langle L_{\tau_1} X_o(\tau_1,\omega)\, (L_{\tau_2} X_o(\tau_2,\omega))^* \rangle$$

$$= \langle L_{\tau_1} L_{\tau_2}^* X_o(\tau_1,\omega)\, X_o^*(\tau_2,\omega) \rangle$$

$$= \langle L_{\tau_1} L_{\tau_2}^* \langle X_o(\tau,\omega)\, X_o^*(\tau_2,\omega) \rangle$$

$$= L_{\tau_1} L_{\tau_2}^* R_{X_o X_o}(\tau_1,\tau_2) , \qquad (2.69)$$

where the subscript on the operator denotes that the operator acts on the stochastic process with the same argument as the subscript. To illustrate application of (2.69), simple linear differential and integral operators are considered. First, let L be $\frac{d}{dt}(\cdot)$. Then by (2.67b)

$$\mu_Y(t) = \frac{d}{dt}\mu_X(t) , \qquad (2.70a)$$

and by (2.69),

$$R_{Y_o Y_o}(t_1,t_2) = \frac{\partial^2 R_{X_o X_o}(t_1,t_2)}{\partial t_1\, \partial t_2} \qquad (2.70b)$$

For higher derivatives, that is,

$$L(\cdot) = \frac{d^n}{dt^n}(\cdot) , \qquad (2.71a)$$

we have

$$\mu_Y(t) = \frac{d^n}{dt^n}\ \mu_X(t) \tag{2.71b}$$

and

$$R_{Y_o Y_o}(t_1,t_2) = \frac{\partial^{2n} R_{X_o X_o}(t_1,t_2)}{\partial t_1^{\,n}\ \partial t_2^{\,n}} \tag{2.71c}$$

All these relations, (2.70) and (2.71), can also be derived directly from the definition of the quadratic mean derivative (2.39).

Another important class of linear operators are integral operators of the following type:

$$Y(t,\omega) = \int_T g(t,\tau)\ X(\tau,\omega)\ d\tau\ , \tag{2.72}$$

where $g(t,\tau)$ is a deterministic function. $g(t,\tau)$ may be a Green's function or the impulse response of a linear system. From (2.67) and (2.69) we have

$$\mu_Y(t) = \int_T g(t,\tau)\ \ \mu_X(\tau)\ d\tau \tag{2.73a}$$

and

$$R_{Y_o Y_o}(t_1,t_2) = \int_T \int_T g(t_1,\tau_1)\ g(t_2,\tau_2)\ \ R_{X_o X_o}(\tau_1,\tau_2)\ d\tau_1\ d\tau_2 \tag{2.73b}$$

The expressions (2.73) can be also derived directly from the definition of the quadratic mean integrals. It is also a special case of the integral (2.40). The results of this section will be frequently used in the subsequent chapters of this dissertation.

The special case of a linear transformation of Gaussian processes is of interest. The characteristic functional of real stochastic processes $X(t,\omega)$ and $Y(t,\omega)$ is

$$g_Y(\lambda) = \langle \exp [\, i \lambda Y(t,\omega) \,] \rangle$$

$$= \langle \exp [\, i \lambda L X(t,\omega) \,] \rangle , \qquad (2.74)$$

where λ is a linear functional. For continuous stochastic processes, λ is an integral

$$\lambda = \int_T \lambda_o(t) \; Y(t,\omega) \, dt , \qquad (2.75)$$

and for discrete stochastic processes, λ is a sum

$$\lambda_N = \sum_{n=1}^{N} \lambda_n(t_n,\omega) \quad . \qquad (2.76)$$

Since λ is a linear functional the relation (2.74) shows that the characteristic functionals for $X(t,\omega)$ and $Y(t,\omega)$ are related by

$$g_Y(\lambda) = g_X [\lambda L] \qquad (2.77)$$

If the real stochastic process $X(t,\omega)$ has a Gaussian distribution, its characteristic functional is (Pugachev, 1965, p. 193):

$$g_X(\lambda) = \exp\left[i \lambda \mu_X(t) - \lambda_{t_1} \lambda_{t_2} C_{XX}(t_1,t_2) \right] . \qquad (2.78)$$

Using either (2.77) or (2.69) and (2.67) the characteristic functional of $Y(t,\omega)$ is given by

$$g_Y(\lambda) = \exp\left[i\lambda L \mu_X(t) - \frac{1}{2} (\lambda L)_{t_2} C_{XX}(t_1,t_2)\right] =$$

$$= \exp\left[i\lambda L \mu_X(t) - \frac{1}{2} \lambda_{t_1} \lambda_{t_2} (L_{\tau_1} L_{\tau_2} C_{XX}(\tau_1,\tau_2)) \right] \qquad (2.79)$$

This relation (2.79) shows that for Gaussian stochastic processes the characteristic functional remains invariant under linear transformation. Since the multivariate distribution functions can be determined from the characteristic functional, it can be concluded that a linear transformation of Gaussian processes yields Gaussian processes. The spectral expansion of a Gaussian process is also a Gaussian process, because spectral expansion is a linear transformation of a random process. As it was stated before, all the multivariate distribution function can be obtained from the characteristic functional. But, from (2.78) it can be seen that the characteristic functional for the Gaussian process is completely determined if the mean and covariance of the process are known. Hence, a Gaussian process is completely determined if its mean and covariance are known. For the same reason a wide-sense stationary Gaussian process is also strictly stationary.

Theorem 2. A second order stochastic process $X(t,\omega)$ is wide-sense stationary if and only if its bifrequency power spectral density $\Phi_{XX}(s_1,s_2)$ is equal to $\delta(s_2-s_1)\ \Phi_{XX}(s_1)$ [or equivalently to $\delta(s_2-s_1)\ \Phi_{XX}(s_2)$].

This result will be used in Chapter IV to identify wide-sense stationary and nonstationary components of a random wave function. This is done by computing the bifrequency spectral density $\Phi_{XX}(s_1,s_2)$ and identifying the coefficients of $\delta(s_2-s_1)$ as the wide-sense stationary parts of the spectral density. The terms which do not contain $\delta(s_2-s_1)$ are identified as nonstationary parts of the spectral density.

2.8 Linear Transformations of Stochastic Processes

In this section some of the known results from linear deterministic operator theory are presented. A linear deterministic system is shown in Figure 2.2. Let the input to the system be $X(t,\omega)$ and the output to the system be $Y(t,\omega)$, $t\epsilon T$ and $\omega\epsilon\Omega$ on the probability space $(\Omega,\mathcal{Z},P)$. Mathematically the relation between the input and output is

$$Y(t,\omega) = L\ X(t,\omega), \tag{2.65}$$

where L is a linear operator. [An operator L is said to be linear if

$$L\ [a\ x_1(t) + b\ x_2(t)] = a\ L\ x_1(t) + b\ L\ x_2(t) \tag{2.66}$$

for every constant a and b, and for every function $x_1(t)$ and $x_2(t)$.] It is easy to compute the expectation and the covariance

2.9 Definitions and Theorems from the Theory of Random Operator Equations

Probabilistic functional analysis, developed mainly by Špaček (1955), Hanš (1961), and Bharucha-Reid (1960, 1964, 1965) provides rigorous definitions and useful existence theorems for random operator equations. The main results of this work have been conveniently summarized by Hanš (1961). Some of the definitions and theorems that are used later in this dissertation are extracted from his work.

Let $(\Omega,\mathfrak{A},P)$ denote a probability space with a complete probability measure P; that is, Ω is a non-empty set, $\mathfrak{A}$ is a σ-algebra of subsets of the space Ω, and $P(A)$ is a probability measure defined on the sets A of $\mathfrak{A}$.

In this section, X and Z are arbitrary separable Banach spaces, $\mathfrak{X}$ and $\mathfrak{b}$ the σ-algebras of all Borel subsets of the spaces X and Z, respectively (Zaanen, 1953).

Next, the concepts of "generalized random variable" and "random transformation" are defined.

A mapping V of the space Ω into the space Z is called a <u>generalized random variable</u> if $\{\omega: V(\omega) \in B\} \in \mathfrak{A}$ for all $B \in \mathfrak{b}$.

Two generalized random variables $V(\omega)$ and $W(\omega)$ are said to be <u>equivalent</u> if $V(\omega) = W(\omega)$ with probability one.

A mapping T of the Cartesian product space $\Omega \times X$ into the space Z is called a random transformation if $T(\cdot,x)$ is, for every $x \in X$, a generalized random variable.

In the most general form, a random operator equation is written as

$$T[\omega, x(\omega)] = z(\omega) \quad , \tag{2.80}$$

where T is a random transformation of the Cartesian product space $\Omega \times X$ into the space Z. z is a generalized random variable with values in the space Z. As is seen from (2.80), the solution of a random operator equation does, in general, depend on the choice of $\omega \in \Omega$. If the solution of (2.80) satisfies the measurability condition (i.e., it is a random variable*) we call it a random solution of the operator equation. Hence, the following definition:

Every generalized random variable x with values in the space X satisfying the condition $T[\omega, x(\omega)] = z(\omega)$ with probability one will be called the <u>random solution</u> of the random operator equation (2.80).

The following theorems which are useful for our investigation have been stated and proved by Hanš (1961).

<u>Theorem 3</u>. Let T be an almost surely linear bounded random transformation of the Cartesian product space $\Omega \times X$ into the space X. Then for every real number $\lambda \neq 0$ such that

$$P\left(\bigcup_{n=1}^{\infty} \{\omega: \| T^n(\omega,\cdot) \| \leq |\lambda|^n\}\right) = 1 , \qquad (2.81)$$

there exists a linear bounded random transformation S that is the inverse of the random transformation $(T - \lambda I)$ and it satisfied

$$P\left(\bigcap_{x \in X} \{\omega: S(\omega,x) = (-1/\lambda) \sum_{n=1}^{\infty} \lambda^{-n}(\omega,x)\}\right) = 1 \qquad (2.82)$$

where the sum converges uniformly. I denotes the identity operator, and $\| \ \|$ denotes the norm in the Banach space (Zaanen, 1953).

*Discussed in section 2.2.

In many applications, it is sometimes convenient to use a different version of the above theorem.

Theorem 4. Let T be a random transformation of the Cartesian product space $\Omega \times X$ into space X which is, for every $\omega \in \Omega$, linear and bounded. Then, for every real number $\lambda \neq 0$, the set

$$\Omega(\lambda) = \{\omega : \|T(\omega,\cdot)\| < |\lambda|\} \tag{2.83}$$

belongs to the σ-algebra $\mathfrak{A}$, the random transformation $(T-\lambda I)$ is invertible for every $\omega \in \Omega(\lambda)$, the resolvent operator $\Gamma(\omega,\lambda,\cdot)$ exists for every $\omega \in \Omega(\lambda)$ and, for these ω, the resolvent operator is given by

$$\Gamma(\omega,\lambda,\cdot) = -\Sigma_{n=1}^{\infty} \lambda^{-n} T^{n-1}(\omega,\cdot) \quad . \tag{2.84}$$

Furthermore, for every $\omega \in \Omega(\lambda)$ the solution $s(\omega)$ of the operator equation

$$T(\omega,\zeta) - \lambda\zeta = z(\omega) \tag{2.85}$$

is, for every generalized random variable z with values in the space X, given by

$$s(\omega) = \Gamma[\omega,\lambda,z(\omega)] \tag{2.86}$$

where the resolvent operator $\Gamma(\omega,\lambda,\cdot)$ and, consequently also the solution s, are measurable with respect to the σ-algebra $\Omega(\lambda) \cap \mathfrak{A}$.

Next, three theorems which are useful in establishing the existence of solutions of the random integral equations are stated. Let C denote the space of all continuous functions defined on the closed interval $[0,d]$, $0 < d$.

If the norm x is given by

$$\| x \| = \max_u |x(u)| \quad ; \quad 0 \leq u \leq d, \tag{2.87}$$

the space C becomes a separable Banach space.

First, a theorem that gives the relationship between the measurability of the random integral operator and the measurability of its kernel is stated. Let K denote the space of all functions k defined and bounded on the set $[0,d] \times [0,d]$ all of whose discontinuity points are located on a finite number of curves $v = \phi_i(u)$ and such that for every $u \in [0,d]$, $v \in [0,d]$, and for every sequence of real numbers $d \geq \delta_1 > \delta_2 > \cdots > \delta_n \to 0$, we have $k(u,0) = \lim_{n\to\infty} k(u,\delta_n)$ and $k(u,v) = \lim_{n\to\infty} k(u,v-\delta_n)$, provided $\delta_1 \leq v$ in the latter case. If the norm of k is $\| k \| = \sup |k(u,v)|$, where sup is taken over $u \in [0,d]$ and $v \in [0,d]$, then it can be shown that K is a separable normed linear space.

Theorem 5. Let k be a mapping of the space Ω into the space K and let the mapping T of the Cartesian product space $\Omega \times C$ into the space C be defined for every $\omega \in \Omega$ by

$$T(\omega,x) = \int_0^d k(\omega,\cdot,u)\, x(u)\, du \quad . \tag{2.88}$$

Then, the mapping T is for every $\omega \in \Omega$ a compact linear transformation of the space C into itself. Furthermore, T is a random transformation.

The next theorem gives a sufficient condition for the invertibility of a linear random transformation.

Theorem 6. Let all the assumptions of Theorem 5 be fulfilled. Let, in addition, the real number λ satisfy the inequality $d\,\|k(\omega,\cdot,\cdot)\| < |\lambda|$ with probability one. Then the linear random transformation $(T - \lambda I)$ is invertible; that is,

$$P\{\omega:(\omega,\lambda) \in \rho(T)\} = 1, \tag{2.89}$$

where $\rho(T)$ denotes the set of those pairs $(\omega,\lambda) \in \Omega \times R$ for which the linear random transformation $(T - \lambda I)$ has a linear bounded inverse. R denotes the real number axis.

The above theorem applies to random Fredholm integral equations. The next theorem is its analogue for the Volterra kernels.

Theorem 7. Let all the assumptions of Theorem 5 be fulfilled. Let, in addition, the kernel k satisfy the condition $P\{\omega: k(\omega,u,v) = 0\} = 1$ for every $0 \le u < v \le d$. Then, for every real number $\lambda \neq 0$ the linear random transformation $(T - \lambda I)$ is invertible.

Proof. Similarly to the classical proof of the convergence of Volterra integral equation*, it follows that for almost every $\omega \in \Omega$ and every $n = 1,2,\ldots$, we have

$$\|T^n(\omega,\cdot)\| \le |\lambda d|^n \, \|k(\omega,\cdot,\cdot)\|^n / n\,! \tag{2.90}$$

Hence, Theorem 3 is applicable and the Neumann series expansion (2.81) converges uniformly.

The above theorems give sufficient conditions for the existence, uniqueness, and measurability of random solution of the random

---

*Zaanen (1953)

operator equations. These are sufficient conditions; that is, solutions may also exist under weaker conditions. Unfortunately, no necessary and sufficient conditions are known for the existence of a solution of the general random operator equations. The main restriction in all the above theorems in that T is almost surely linear and bounded.

Many interesting stochastic processes are not bounded, hence the preceding theorems are not strictly applicable. We can use a truncation method, suggested by Loève (1963), to overcome this difficulty. The random variable $X(\omega)$ is <u>truncated</u> at c when we replace $X(\omega)$ by $X^c(\omega)$, where

$$X^c(\omega) = X(\omega) \text{ if } |X(\omega)| < c,$$

$$X^c(\omega) = 0 \text{ if } |X(\omega)| \geq c \quad ,$$

and

$$c > 0 \text{ and finite.} \tag{2.91}$$

Then, all the moments of $X^c(\omega)$ exist and are finite. We can always select c sufficiently large so as to make $P[X(\omega) \neq X^c(\omega)] = P[|X(\omega)| \geq c]$ arbitrarily small. Whenever boundedness of the stochastic processes is needed in the application of the preceding theorems (Theorems 3 through 7), it is assumed that the stochastic processes are truncated in the above sense.

CHAPTER III

GENERAL METHODS OF SOLUTION

3.1 Introduction

In this chapter, two methods for solving differential equations with randomly time-varying coefficients are developed. In the first case, we consider a stochastic differential equation of the form

$$\mathcal{L}\ y(t,\omega) = x(t,\omega) \quad ,$$

where

$$\mathcal{L} = \sum_{\nu=0}^{n} a_{\nu}(t,\omega)\ \frac{d^{\nu}}{dt^{\nu}} \quad ,$$

and $x(t,\omega)$ and the $a_{\nu}(t,\omega)$ are random functions whose statistics are known and defined on $t\epsilon T$, $\omega\epsilon\Omega$ on $(\Omega,\mathfrak{A},P)$. $(\Omega,\mathfrak{A},P)$ denotes a probability measure space; i.e., Ω is a non-empty abstract set, $\mathfrak{A}$ is a σ algebra of subsets of Ω, and P is a complete probability measure on $\mathfrak{A}$. It is further assumed that the random operator $\mathcal{L}$ is the sum of an invertible operator L and a random operator R. The objective is to determine a stochastic Green's function (Adomian, 1964) for the stochastic operator $\mathcal{L}$ in terms of the deterministic Green's function for L and the appropriate "statistical measures" of the forcing function $x(t,\omega)$ and the randomly time-varying coefficients $a(t,\omega)$. The term "statistical measures" is used as a general term for the quantities that characterize stochastic processes. For example, expectations or means, spectral

densities and correlation functions are statistical measures. The integral kernel which expresses the desired statistical measure of the output process in terms of the corresponding statistical measures of the input and appropriate statistical measures of the stochastic coefficients is called the "stochastic Green's function". Adomian (1967) has developed an iterative process for finding the stochastic Green's functions for the expectation and the autocorrelation function of the output process. This iterative method achieves the desired seperations of ensemble averages without a priori restrictions to perturbation-type approaches or recourse to the generally non-valid closure approximations of the hierarchy equation methods.

The similarity between Adomian's iterative method and the solution of an integral equation by means of Neumann series suggests that the stochastic Green's function can be expressed in terms of the deterministic Green's function for the operator L and the resolvent kernel for the Volterra integral equation. In this chapter, the ensemble average and correlation function of the dependent variable are expressed in terms of the Green's function of the operator L, the average and the correlation function of the forcing function $x(t,\omega)$ and the resolvent kernel of the Volterra integral equation. The resolvent kernel is a function of appropriate statistical measures of the random coefficients and the Green's function of the deterministic operator. By finding the resolvent kernel, the solution of the problem of finding the average and the correlation function of the dependent variable is expressed for a class of forcing functions. The Neumann series expansion is exceedingly laborious; for this reason, other methods for solving the

problem are investigated. The solution of the stochastic integral equation is simplified if the kernel of the integral equation is degenerate. In many cases, the problem of solving a stochastic differential equation, with a time invariant deterministic part, reduces to a problem of solving an integral equation with a degenerate kernel. Three different cases of degenerate kernels are considered. In the first and simplest case, the degenerate kernel consists of a single term. In the second case, the degenerate kernel is a sum of n products. In the third case, the degenerate kernel is a product of the state transition matrix and a matrix of stochastic coefficients. The last case is applicable to the state space formulation of control system problems where the system matrix is the sum of a deterministic time-invariant matrix and a stochastic coefficient matrix.

3.2 Integral Equation Formulation and Solution by Neumann Series Expansion

Following Adomian's (1967) approach, let $\mathcal{L}$ be an n^{th} order stochastic differential operator, such that

$$\mathcal{L} = \sum_{\nu=0}^{n} a(t,\omega) \frac{d^{\nu}}{dt^{\nu}} \quad , \tag{3.1a}$$

where $a_{\nu}(t,\omega)$ are random functions, $t\epsilon T$ and $\omega\epsilon\Omega$ $(\Omega,\mathcal{X},P)$. Let the operator $\mathcal{L}$ be separable into a deterministic operator L and a random operator R. In particular, let the random coefficients of $\mathcal{L}$ be of the form

$$a_\nu(t,\omega) = \beta_\nu(t) + \alpha_\nu(t,\omega) \ , \qquad (3.1b)$$

where $\beta_\nu(t)$ are deterministic functions of time and $\alpha_\nu(t,\omega)$ are stochastic processes. $\beta_\nu(t)$ can be either the ensemble average of $a_\nu(t,\omega)$ or some other convenient function of time. For example, it may be possible to choose $\beta_\nu(t)$ so that the inversion of the deterministic differential operator L is simplified. It is assumed that L is an invertible differential operator; that is, the Green's function $G(t,\tau)$ for the differential operator L is known or can be constructed. One can associate with the operator $\mathcal{L}$ a stochastic differential equation

$$\mathcal{L}\, y(t,\omega) = x(t,\omega) \qquad t\epsilon T \text{ and } \omega\epsilon\Omega \text{ on } (\Omega,\mathcal{Z},P). \qquad (3.2)$$

The forcing function $x(t,\omega)$ can be either deterministic or random. For greater generality, let it be random. It is also assumed that $x(t,\omega)$ is statistically independent of the random coefficients. By the assumption that $\mathcal{L} = L + R$, the equation (3.2) can be written as follows:

$$L\, y(t,\omega) = x(t,\omega) - R\, y(t,\omega) \qquad (3.3a)$$

where

$$L = \sum_{\nu=0}^{n} \beta_\nu(t) \frac{d^\nu}{dt^\nu} \qquad (3.3b)$$

and

$$R = \sum_{\nu=0}^{n} a_\nu(t,\omega) \frac{d^\nu}{dt^\nu} \qquad (3.3c)$$

By the assumption that the Green's function for the deterministic operator L is known, (3.3) can be converted into an integral equation:

$$y(t,\omega) = L^{-1} x(\tau,\omega) - L^{-1} R\, y(\tau,\omega) \quad , \qquad (3.4a)$$

where,

$$L^{-1} = \int_0^t d\tau\, G(t,\tau) \qquad (3.4b)$$

and

$$L^{-1} R = \int_0^t d\tau\, G(t,\tau) \sum_{\nu=0}^{n} a_\nu(\tau,\omega) \frac{d^\nu}{d\tau^\nu} \quad . \qquad (3.4c)$$

The use of 0 as the lower limit assumes that either the system was initially at rest or the initial conditions have been taken into account in the construction of the Green's function $G(t,\tau)$. The upper limit is t for causal systems since, for causal systems, $G(t,\tau) =$ for $t < \tau$. To simplify the notation, let

$$L^{-1} x(t,\omega) = F(t,\omega), \qquad (3.5a)$$

and let

$$K(t,\tau) = G(t,\tau) \sum_{\nu=0}^{n} \alpha_{\nu}(\tau,\omega) \frac{d^{\nu}}{d\tau^{\nu}} \quad . \tag{3.5b}$$

With this notation, Equation (3.4) can be expressed as an integral equation which resembles the Volterra integral equation:

$$y(t,\omega) = F(t,\omega) + \lambda \int_{o}^{t} K(t,\tau)\, y(\tau,\omega)\, d\tau \quad . \tag{3.6}$$

In this case, $\lambda = -1$. The solution to Equation (3.6) can be expressed analogously to the Neumann series solution in terms of iterated kernels (Courant and Hilbert, 1953; and Hildebrand, 1952):

$$y(t,\omega) = F(t,\omega) + \sum_{m=1}^{\infty} \lambda^{m} \int_{o}^{t} K_{m}(t,\tau)\, F(\tau,\omega)\, d\tau \quad , \tag{3.7}$$

where $K_{m}(t,\tau)$ is defined by the recurrence formula

$$K_{m}(t,\tau) = \int_{o}^{t} K(t,\tau_{1})\, K_{m-1}(\tau_{1},\tau)\, d\tau_{1} \tag{3.8a}$$

and

$$K_{1}(t,\tau) = K(t,\tau) \quad . \tag{3.8b}$$

If the sum converges uniformly,* the order of summation and integration can be interchanged in (3.7), which then becomes:

$$y(t,\omega) = F(t,\omega) + \lambda \int_0^t \left[\sum_{m=0}^{\infty} \lambda^m K_{m+1}(t,\tau)\, F(\tau,\omega)\, d\tau \right] . \tag{3.9}$$

Then equation (3.9) can be written in terms of the resolvent kernel $\Gamma(t,\tau,\omega)$ as follows:

$$y(t,\omega) = F(t,\omega) + \lambda \int_0^t \Gamma(t,\tau,\lambda,\omega)\, F(\tau,\omega)\, d\tau \quad , \tag{3.10a}$$

where

$$\Gamma(t,\tau,\lambda,\omega) \sum_{m=0}^{\infty} \lambda^m K_{m+1}(t,\tau,\lambda,\omega). \tag{3.10b}$$

Since, in this case, λ is -1, the resolvent kernel is written from now on without λ in its argument. At this point, we have nothing essentially different from classical differential equation theory. The difficulty arises because $y(t,\omega)$ is not a physically significant quantity, only its statistical measures are. The mean value of the dependent variable $y(t,\omega)$ can be computed by taking the ensemble

*Discussed later in this section.

average of equation (3.10). It must be noted that because of the assumption that the random coefficients and forcing function are independent, the problem simplifies in that the resolvent kernel $\Gamma(t,\tau,\omega)$ and $F(\tau,\omega)$ are statistically independent of one another. Hence, the ensemble averages separate and the ensemble average of $y(t,\omega)$ is given by:

$$< y(t,\omega) = < F(t,\omega) > - \int_0^t < \Gamma(t,\tau,\omega) > < F(\tau,\omega) > d\tau \ , \tag{3.11a}$$

where from (3.4b) and (3.5a)

$$< F(t,\omega) > = \int_0^t G(t,\tau) < x(t,\omega) > d\tau \ . \tag{3.11b}$$

In this case, the stochastic Green's function is simply $< \Gamma(t,\tau,\omega) >$.

Similarly, the autocorrelation function of $y(t,\omega)$ can be computed. For greater generality, $y(t,\omega)$ is taken to be a complex function. Then, the autocorrelation function of $y(t,\omega)$ denoted by $R_{yy}(t_1,t_2)$ is:

$$R_{yy}(t_1,t_2) = < y(t_1,\omega)\ y^*(t_2,\omega) >$$

$$= < \left\{ F(t_1,\omega) - \int_0^{t_1} \Gamma(t_1,\tau_1,\omega)\ F(\tau_1,\omega)\ d\tau_1 \right\}$$

$$\left\{ F^*(t_2,\omega) - \int_0^{t_2} \Gamma^*(t_2,\tau_2,\omega)\ F^*(\tau_2,\omega)\ d\tau_2 \right\} >$$

$$= \langle F(t_1,\omega)\, F^*(t_2,\omega)\rangle$$

$$- \int_0^{t_1} \langle \Gamma(t_1,\tau_1,\omega)\rangle \langle F(\tau_1,\omega)\, F^*(t_2,\omega)\rangle\, d\tau_1$$

$$- \int_0^{t_2} \langle \Gamma^*(t_2,\tau_2,\omega)\rangle \langle F(t_1,\omega)\, F^*(\tau_2,\omega)\rangle\, d\tau_2$$

$$+ \int_0^{t_1}\int_0^{t_2} \langle \Gamma(t_1,\tau_1,\omega)\, \Gamma^*(t_2,\tau_2,\omega)\rangle \langle F(\tau_1,\omega)\, F^*(\tau_2,\omega)\rangle\, d\tau_1\, d\tau_2 \,. \qquad (3.12a)$$

The quantity $\langle F(t_1,\omega)\, F^*(t_2,\omega)\rangle$ can be denoted by $R_{FF}(t_1,t_2)$ and (3.12a) can be written more compactly as:

$$R_{yy}(t_1,t_2) = R_{FF}(t_1,t_2)$$

$$- \int_0^{t_1} \langle \Gamma(t_1,\tau_1,\omega)\rangle R_{FF}(\tau_1,t_2)\, d\tau_1$$

$$- \int_0^{t_2} \langle \Gamma^*(t_2,\tau_2,\omega)\rangle R_{FF}(t_1,\tau_2)\, d\tau_2$$

$$+ \int_0^{t_1}\int_0^{t_2} \langle \Gamma(t_1,\tau_1,\omega)\, \Gamma^*(t_2,\tau_2,\omega)\rangle R_{FF}(\tau_1,\tau_2)\, d\tau_1\, d\tau_2 \,, \qquad (3.12b)$$

where

$$R_{FF}(t_1,t_2) = \int_0^{t_1}\int_0^{t_2} G(t_1,\tau_1)G^*(t_2,\tau_2) < x(\tau_1,\omega)\, x^*(\tau_2,\omega) > d\tau_1\, d\tau_2$$

$$= \int_0^{t_1}\int_0^{t_2} G(t_1,\tau_1)G^*(t_2,\tau_2)\, R_{xx}(\tau_1,\tau_2)\, d\tau_1\, d\tau_2 \quad . \tag{3.12c}$$

Equation (3.12) looks quite complicated, yet it expresses one of the most important properties of y. It is unfortunate that many of the interesting properties of y are given by nonlinear functions. If the coefficients of the differential operator are strictly deterministic, then the last terms in (3.12b) vanish and $R_{yy}(t_1,t_2)$ is given by (3.12c) which agrees with the well known result for non-random systems. The cross correlation between the output and input can be also found by using equation (3.10):

$$R_{yx}(t_1,t_2) = < y(t_1,\omega)\, x^*(t_2,\omega) >$$

$$= \int_0^{t_1} G(t_1,\tau_1)\, R_{xx}(\tau_1,t_2)\, d\tau_1$$

$$- \int_0^{t_1} < \Gamma(t_1,\tau_1,\omega) > \left(\int_0^{\tau_1} G(\tau_1,\tau_0)\, R_{xx}(\tau_0,t_2)\, d\tau_0 \right) d\tau_1 \tag{3.13}$$

Examination of equations (3.11), (3.12b) and (3.13) sheds some light on the stochastic Green's functions (Adomian, 1964) for a differential operator which can be separated into a sum of the deterministic operator L and the random operator R. In all three cases, the first term in the expressions is due to the deterministic operator alone. If the randomness of the coefficients of Γ vanishes, only the first terms remain. The succeeding terms are complicated functions of random coefficients and deterministic Green's function for the operator L. When higher order coefficients ($\nu \geq 1$) are random, construction of the resolvent kernel involves differentiation as indicated by (3.5b). This is not too surprising for a relatively complicated problem. In other physical problems, similar complications arise. For example, an expression for the dyadic Green's function associated with the solution of a vector wave equation involves differentiation (Levine and Schwinger, 1951).

In the actual solution of problems, construction of the resolvent kernel presents major difficulties. As shown by Adomian (1967), Neumann series type of iteration can be successfully used when the lowest order coefficient is stochastic. This type of iteration has the advantage that the previously computed term is used in the next term and so on. One can stop the iteration at any time when the remainder term becomes smaller than a prescribed value.

When only the lowest order coefficient a_0 is a random function, it is easy to show the convergence of the Neumann series expansion (3.7). Sufficient conditions for the uniform convergence of this Neumann series expansion are simply the hypotheses of Theorem 5 and 7

of section 2.9. The hypotheses of Theorem 5 require that $G(t,\tau)\ \alpha_o(\tau,\omega)$ is a mapping of the space Ω into space K and that the mapping

$$\int_0^t d\tau\ G(t,\tau)\ \alpha_o(\tau,\omega)\ F(\tau,\omega)$$

of the Cartesian product space Ω x C into the space C is defined for every $\omega\epsilon\Omega$ and every $F(\tau,\omega)\ \epsilon\ C$. C denotes the space of all continuous functions defined on the closed interval $[0,d]$, $0 < d$. The space K has been defined in section 2.9. The additional hypotheses of Theorem 7 simply state that the kernel of the integral equation is a Volterra kernel with probability one. Under these conditions, the integral equation (3.6) is invertible and the Neumann series expansion converges uniformly.

When the higher order coefficients are also stochastic, it is more difficult to carry out the Neumann series expansion and to establish the sufficient conditions for the convergence of the series. It can be seen from (3.5) that the expansion as it stands requires differentiation in addition to integration. For computation and for investigation of convergence, it is convenient to express the n^{th} order differential equation as n first-order differential equations which are in the matrix form:

$$\dot{\underline{y}}(t,\omega) = \underline{a}(t,\omega)\ \underline{y}(t,\omega) + \underline{U}\ \underline{x}(t,\omega) \quad , \tag{3.14}$$

where $\underline{\dot{y}(t,\omega)}$ and $\underline{y(t,\omega)}$ are n-dimensional vectors, $\underline{a(t,\omega)}$ is an n x n coefficient matrix, $\underline{U}$ is an n x r matrix, and $\underline{x(t,\omega)}$ is an r-dimensional vector. This allows the system to have r different inputs. In the single forcing function case, r is equal to one. In control system theory, this matrix formulation is called state space representation. The methods of converting the single n^{th} order differential equation into state space representation are discussed in many texts on control system theory and need no elaboration here (Schultz and Melsa, 1967; DeRusso, Roy and Close, 1965; and Lapidus and Luus, 1967). Because of the original assumption that the differential operator can be expressed as a sum of a deterministic operator and a stochastic operator, we assume that the coefficient matrix $\underline{a(t,\omega)}$ can be expressed as

$$\underline{a(t,\omega)} = \underline{\beta(t)} + \underline{\alpha(t,\omega)} \quad , \tag{3.15}$$

where $\underline{\beta(t)}$ is a matrix with deterministic elements and $\underline{\alpha(t,\omega)}$ is a matrix with stochastic elements. Using (3.15), the differential equation (3.14) can again be written in the integral equation form

$$\underline{y(t,\omega)} = \underline{F(t,\omega)} + \lambda \int_0^t \underline{\Phi(t,\tau)}\ \underline{\alpha(\tau,\omega)}\ \underline{y(\tau,\omega)}\ d\tau, \tag{3.16a}$$

where

$$\int_0^t d\tau\ \underline{\Phi(t,\tau)} \tag{3.16b}$$

is the inverse operator of the differential equation

$$\dot{\underline{y}}(t,\omega) = \underline{\beta}(t) \quad \underline{y}(t,\omega) \quad . \tag{3.17}$$

The term $\underline{F}(t,\omega)$ is obtained by applying the inverse operator (3.16b) to all forcing terms and non-zero initial conditions. In control system terminology, $\underline{\Phi}(t,\tau)$ is called the state transition matrix. In this case, λ is equal to one. The Neumann series expansion of (3.16a) gives the following iterative solution:

$$\underline{y}(t,\omega) = \sum_{n=0}^{\infty} \lambda^n \quad \underline{y}_n(t,\omega)$$

$$\underline{y}_o(t,\omega) = \underline{F}(t,\omega)$$

$$\underline{y}_1(t,\omega) = \int_o^t \underline{\Phi}(t,\tau) \; \underline{\alpha}(\tau,\omega) \; \underline{y}_o(\tau,\omega) \; d\tau$$

$$= \int_o^t \underline{\Phi}(t,\tau) \; \underline{\alpha}(\tau,\omega) \; \underline{F}(\tau,\omega) \; d\tau \quad ,$$

(3.18)

or, in general,

$$\underline{y}_n(t,\omega) = \int_o^t \underline{\Phi}(t,\tau) \; \underline{\alpha}(\tau,\omega) \; \underline{y}_{n-1}(\tau,\omega) \; d\tau \quad .$$

As before, the resolvent kernel for the integral equation (3.16) can be written in terms of iterated kernels by

$$\underline{\Gamma(t,\tau,\lambda,\omega)} = \sum_{n=0}^{\infty} \lambda^n \underline{K_{n+1}(t,\tau,\omega)} \quad ,$$

where the iterated kernels are defined by the recurrence formula

$$\underline{K_m(t,\tau,\omega)} = \int_0^t \underline{K(t,\tau_1)} \; \underline{K_{m-1}(\tau_1,\tau)} \; d\tau_1$$

and

$$\underline{K_1(t,\tau,\omega)} = \underline{K(t,\tau,\omega)} = \underline{\Phi(t,\tau)} \; \underline{\alpha(\tau,\omega)} \quad . \tag{3.19}$$

The solution to the matrix stochastic differential equation (3.14) can be written in terms of the matrix resolvent kernel as

$$\underline{y(t,\omega)} = \underline{F(t,\omega)} + \lambda \int_0^t \underline{\Gamma(t,\tau,\lambda,\omega)} \; \underline{F(\tau,\omega)} \; d\tau \quad . \tag{3.20}$$

This expression is completely analogous to (3.10). Construction of the resolvent kernel in (3.10) requires differentiation in addition to integration, whereas, in this case, the resolvent kernel can be constructed by iterated integration and matrix multiplication. The previously computed terms are used to compute the next term and so on.

The ensemble average of $y(t,\omega)$ is

$$< \underline{y(t,\omega)} > = < \underline{F(t,\omega)} > + \int_0^t < \underline{\Gamma(t,\tau,\omega)} > < \underline{F(\tau,\omega)} > d\tau \quad .$$

(3.21)

The stochastic Green's function for the ensemble average of $y(t,\omega)$ is simply $\langle \Gamma(t,\tau,\omega) \rangle$. The expression for the covariance matrix is

$$\langle \underline{y(t_1,\omega)} \ \underline{y^\dagger(t_2,\omega)} \rangle = \langle \underline{F(t_1,\omega)} \ \underline{F^\dagger(t_2,\omega)} \rangle$$

$$+ \int_0^{t_2} \langle \underline{F(t_1,\omega)} \ \underline{F^\dagger(\tau,\omega)} \rangle \langle \underline{\Gamma^\dagger(t_2,\tau,\omega)} \rangle \, d\tau$$

$$+ \int_0^{t_2} \langle \underline{\Gamma(t_1,\tau,\omega)} \rangle \langle \underline{F(\tau,\omega)} \ \underline{F^\dagger(t_2,\omega)} \rangle \, d\tau$$

$$+ \int_0^{t_1} \int_0^{t_2} \langle \underline{\Gamma(t_1,\tau,\omega)} \ \underline{F(\tau,\omega)} \ \underline{F^\dagger(\sigma,\omega)} \ \underline{\Gamma^\dagger(t_2,\sigma,\omega)} \rangle d\tau d\sigma \ , \tag{3.22}$$

where † denotes complex conjugate transpose (hermetian conjugate) of the matrix. The ensemble averages of (3.22) also separate in the last term because $\underline{\Gamma(t,\tau,\omega)}$ and $\underline{F(\tau,\omega)}$ are statistically independent, but this cannot be shown with the matrix notation because the order of matrix multiplication must be preserved. Separation of ensemble averages takes place after the matrix multiplication has been carried out.

A sufficient condition for the convergence of the Neumann series (3.19) can be given by a theorem analogous to Theorem 4 of section 2.9.

Theorem 1. Let $\underline{\Phi(t,\tau)}\ \underline{\alpha(\tau,\omega)} = k$ be a mapping of the space Ω into the space* K and let the mapping T of the Cartesian product space $\Omega \times C$ into the space C be defined for every $\omega\epsilon\Omega$ and every $F(\tau,\omega) \epsilon C$ by

$$T(\omega,t,\underline{F(\tau,\omega)}) = \int_0^t \underline{\Phi(t,\tau)}\ \underline{\alpha(\tau,\omega)}\ \underline{F(\tau,\omega)}\ d\tau \ . \tag{3.23}$$

In addition, let the kernel $\underline{\Phi(t,\tau)}\ \underline{\alpha(\tau,\omega)}$ satisfy the condition

$$P\{\omega: \underline{\Phi(t,\tau)}\ \underline{\alpha(\tau,\omega)} = 0\} = 1 \tag{3.24}$$

for every $0 \leq t < \tau \leq d$, then the integral equation

$$\underline{y(t,\omega)} = \underline{F(t,\omega)} + \lambda \int_0^t \underline{\Phi(t,\tau)}\ \underline{\alpha(\tau,\omega)}\ \underline{y(\tau,\omega)}\ d\tau$$

is invertible for every finite λ.

Proof. Consider the norms of the iterated solution:

$$\| \underline{y_o(t,\omega)} \| \leq \| \underline{F(t,\cdot)} \|$$

where $\| \ \|$ denotes any suitable norm for the matrices in the Banach space.

$$\| \underline{y_o(t,\omega)} \| = \| \int_0^t \underline{\Phi(t,\tau)}\ \underline{\alpha(\tau,\omega)}\ \underline{y_o(\tau,\omega)}\ d\tau \|$$

*The space K has been defined in section 2.9.

$$\leq \int_0^t \| \underline{\Phi(t,\tau)} \ \underline{\alpha(\tau,\omega)} \| \ \| \underline{F(\tau,\omega)} \| \ d\tau$$

$$< t \ \| \underline{\Phi(t,\tau)} \ \underline{\alpha(\tau,\omega)} \| \ \| \underline{F(\tau,\omega)} \| \ .$$

The last step follows because the norms $\| \underline{\Phi(t,\tau)} \ \underline{\alpha(\tau,\omega)} \|$ and $\| \underline{F(\tau,\omega)} \|$ are real non-negative numbers.

$$\| \underline{y_2(t,\omega)} \| = \| \int_0^t \underline{\Phi(t,\tau)} \ \underline{\alpha(\tau,\omega)} \ \underline{y_1(\tau,\omega)} \ d\tau \|$$

$$\leq \int_0^t \tau \ \| \underline{\Phi(t,\tau)} \ \underline{\alpha(\tau,\omega)} \|^2 \ \| \underline{F(\tau,\omega)} \| \ d\tau$$

$$\leq \frac{t^2}{2!} \ \| \underline{\Phi(t,\tau)} \ \underline{\alpha(\tau,\omega)} \|^2 \ \| \underline{F(\tau,\omega)} \|$$

or, in general,

$$\| \underline{y_n(t,\omega)} \| \leq \frac{t^n}{2!} \ \| \underline{\Phi(t,\tau)} \ \underline{\alpha(\tau,\omega)} \|^n \ \| \underline{F(\tau,\omega)} \| \tag{3.25}$$

or the norm of the n^{th} transformation T^n satisfies the following inequality

$$\| \underline{T^n} \| \leq \frac{t^n}{n!} \ \| \underline{\Phi(t,\tau)} \ \underline{\alpha(\tau,\omega)} \|^n \ . \tag{3.26}$$

Hence, by Theorem 3 of section 2.9, the integral equation is invertible and its Neumann series expansion converges uniformly for all finite t and λ. In this case, $\lambda = 1$.

Sufficient conditions have been given for the invertibility of the stochastic integral equation (3.16) and for the convergence of the Neumann series expansion. These conditions state that the solution exists for almost all sample functions, but the theorems say nothing about the existence of expected value or covariance of the solution. The sample functions $y(t,\omega)$ may not be second-order stochastic processes. A stochastic process $y(t,\omega)$ is said to be second order if $< | y(t,\omega) |^2 > < \infty$ is satisfied.

Consider now an example which demonstrates that the mean or covariance may not exist. Assume that the probability distribution of the amplitude of the stochastic coefficients $\alpha_\nu(t,\omega)$ is Gaussian (normal). The normality is conserved under repeated differentiation and integration.* Hence, in computing $< y(t,\omega) >$ or $< y(t_1,\omega) y^\dagger(t_2,\omega) >$ by averaging the infinite series term by term, we are computing higher and higher moments of $\alpha_\nu(t,\omega)$. Higher moments of the zero mean Gaussian random variables are related to the second moment, γ^2, by the following expression (Miller, 1964):

$$< \gamma^n > = i^{-n} < \gamma^2 >^{\frac{n}{2}} H_n(0) , \qquad (3.27a)$$

*Discussed in section 2.8.

where $H_n(0)$ is the n^{th} Hermite polynomial given by

$$H_{2n}(0) = \frac{(-1)^n (2n)!}{2^n n!} \qquad n = 0,1,2,\ldots \qquad (3.27b)$$

and

$$H_{2n+1}(0) = 0 \qquad n = 0,1,2,\ldots \qquad (3.27c)$$

Thus the even moments $\to \infty$ as $n \to \infty$. Because of this, the convergence of the averaged series requires further investigation. The terms in the averaged series are

$$\underline{\langle y_o(t,\omega) \rangle} = \underline{\langle F(t,\omega) \rangle}$$

$$\underline{\langle y_1(t,\omega) \rangle} = \int_o^t \langle \underline{\Phi(t,\tau)} \; \underline{\alpha(\tau,\omega)} \rangle \langle \underline{F(\tau,\omega)} \rangle \, d\tau$$

$$= \int_o^t \langle \underline{\gamma(t,\tau,\omega)} \rangle \langle \underline{F(\tau,\omega)} \rangle \, d\tau \quad . \qquad (3.28)$$

The ensemble averages separate in (3.28) because of the statistical independence of $\underline{\alpha(\tau,\omega)}$ and $\underline{F(\tau,\omega)}$. The norm of the general even term is

$$\| < \underline{y_{2k}(t,\omega)} > \| =$$

$$\| \int_0^t \cdots \int_0^{t_{2k-1}} < \underline{\gamma(t,t_1)} \cdots \underline{\gamma(t_{2k-1},t_{2k})} > < \underline{F(t_{2k},\omega)} > dt_1 \cdots dt_{2k} \|$$

$$\leq \frac{t^{2k}}{(2k)!} \| < \underline{\gamma(t,t_1)} \cdots \underline{\gamma(t_{2k-1},t_{2k})} > \| \| < \underline{F(t,\omega)} > \|$$

$$\leq \frac{t^{2k}}{(2k)!} \; n^{2k-1} \; | < M^{2k} > | \; \| < \underline{F(t,\omega)} > \| \quad , \tag{3.29}$$

where we have chosen for the norm of the averaged matrix $< \underline{\gamma} >$ and upper bound of its elements. The odd moments vanish because elements $\underline{\gamma}$ are zero mean Gaussian processes. Recall that $\underline{\gamma}$ is an n x n matrix. For Gaussian random variables, we have:

$$< M^{2k} > = < m^2 >^k \; \frac{(2k)!}{2^k k!} \qquad k = 0,1,2,\ldots$$

and

$$< M^{2k+1} > = 0 \qquad k = 0,1,2,\ldots \quad , \tag{3.30}$$

where $< m^2 >$ is the upper bound of mean squares of the elements.

Hence, the norm of the general term of the Neumann series expansion is:

$$\| < \underline{y_{2k}(t,\omega)} > \| \leq \frac{(nt)^{2k}}{2^k k!} \, n < m^2 >^k \| < \underline{F(t,\omega)} \| . \tag{3.31}$$

The inequality (3.31) means that the averaged Neumann series expansion converges uniformly even if the stochastic coefficient matrix consists of Gaussian random variables. Similar analysis may be applied to the Neumann series expansion of the covariance matrix (3.22). The only term that presents any complications in the analysis is the last term of (3.22).

Inspection of (3.21) and (3.22) shows that for calculation of the second-order statistics of the output process, knowledge of all the moments of the coefficients is required. In many physical problems, such complete knowledge is lacking or it is difficult to obtain. In such cases, one must be satisfied with the approximate solution. However, the iterative mathematical approach suggests that an iterative approach could also be used in constructing a physical model. One could obtain experimental data for computation of the mean and covariance of the stochastic coefficients. Then the first few terms of (3.21) and (3.22) could be used to calculate the approximations for the mean and covariance of $y(t,\omega)$ which could be compared with experimental results. If the agreement between calculated and experimental results is unsatisfactory, more complete data should be obtained on the stochastic coefficients and the procedure should be iterated until the results become adequate.

(If the stochastic coefficients are Gaussian, then the knowledge of their second-order statistics is sufficient for complete solution of the problem.)

In the solution of an actual problem, the Neumann series expansion becomes quite involved. For this reason, other methods of solution are investigated in the following sections. The advantages of the Neumann series expansion are that it is more generally applicable than other methods, sufficient conditions for convergence are known, and it provides useful insight to the nature of stochastic Green's functions.

3.3 Construction of Resolvent Kernel in the Case of a Degenerate Integral Equation

The construction of the resolvent kernel by means of the Neumann series expansion is, in general, laborious. For this reason, it is desirable to investigate other methods of solving the integral equation (3.16). Closed form solutions to the Fredholm and Volterra integral equations can be found if these equations have degenerate kernels (Courant and Hilbert, 1953; Hildebrand, 1952; and Kantorovich and Krylov, 1958). A kernel is said to be degenerate if it can be represented in the form of a finite sum of products of functions of a single (deterministic) variable:

$$K(t,\tau,\omega) = \sum_{k=1}^{n} c_k(t)\, b_k(\tau,\omega) \qquad n < \infty \ . \tag{3.32}$$

Such representation is possible if, for example, the Green's function of the deterministic part of the differential equation (3.2) is a sum of exponentials. The, the terms in (3.32) are:

$$c_k(k) = K_k \exp [s_k t] \qquad (3.33a)$$

and

$$b_k(\tau,\omega) = \exp [-s_k \tau] \alpha_o(\tau,\omega) , \qquad (3.33b)$$

where s_k and K_k are complex numbers. $\alpha_o(\tau,\omega)$ is the random coefficient of the $\nu = 0$ term in (3.1). By considering kernels of the integral equation which can be represented by (3.32) and by (3.33) we have lost generality in two ways:

1) An exponential solution of the form (3.32) and (3.33) implies that the deterministic equation is time-invariant.
2) The form of $b_k(\tau,\omega)$ in (3.33) implies that only $\alpha_o(\tau,\omega)$ is stochastic. This restriction will be removed later when a state function representation of equation (3.2) is considered.

The first restriction is the penalty we pay for the simplification of the computation. The impulse response of most time-invariant linear systems consists of a sum of exponentials (Schultz and Melsa, 1967). Exceptional cases are when the impulse response is either a constant, t or an exponential times t. From the control system point of view, the exponential solutions are the most common. Thus, even

with this specialization, we are still considering an interesting case.

Also, representation of an arbitrary kernel by a degenerate kernel leads to a useful approximation technique. This will be discussed after development of the solution for degenerate kernels.

3.4 Single Term Degenerate Kernel

To demonstrate the essential ideas, we start with the simple case where the degenerate kernel consists of a single term. The integral equation is

$$y(t,\omega) = F(t,\omega) + \lambda \int_0^t K(t,\tau,\omega)\ y(\tau,\omega)\ d\tau\ , \qquad (3.34a)$$

where

$$K(t,\tau,\omega) = c(t)\ b(\tau,\omega)\ . \qquad (3.34b)$$

The functions $c(t)$ and $b(\tau,\omega)$ are

$$c(t) = e^{st}$$

and

$$b(\tau,\omega) = e^{-s\tau}\ \alpha(\tau,\omega)\ , \qquad (3.35)$$

where $\alpha(\tau,\omega)$ is the randomly time-varying coefficient. Note that the factor $c(t)$ may be taken outside the integral in Equation (3.34a) and the remaining integral is some constant independent of t. Thus, $y(t,\omega)$ must be of the form

$$y(t,\omega) = F(t,\omega) + A(t,\omega)\, c(t) \quad , \tag{3.36}$$

where $A(t,\omega)$ is, for the time being, an unknown random process. Substituting Equation (3.36) into Equation (3.34), we obtain:

$$A(t,\omega)\, c(t) = \lambda \int_0^t c(t)\, b(\tau,\omega)\, F(\tau,\omega)\, d\tau + \lambda \int_0^t c(t)\, b(\tau,\omega)\, c(\tau)\, A(t,\omega)\, d\tau \quad . \tag{3.37}$$

Solving for $A(t,\omega)$, we have:

$$A(t,\omega) = \lambda \exp\left[\lambda \int_0^t c(\tau)b(\tau,\omega)\right] \int_0^t \exp[-\lambda c(\sigma)\, b(\sigma,\omega)\, d\sigma]\, F(\tau,\omega)b(\tau,\omega)\, d\tau. \tag{3.38}$$

Substituting Equation (3.38) into Equation (3.36), the solution of the integral equation becomes:

$$y(t,\omega) = F(t,\omega) + c(t)\exp\left[\lambda \int_0^t c(\tau)b(\tau,\omega)d\tau\right] \int_0^t \exp\left[-\lambda \int_0^\tau c(\sigma)b(\sigma,\omega)d\sigma\right] F(\tau,\omega)b(\tau,\omega)\, d\tau. \tag{3.39}$$

$$y(t,\omega) = F(t,\omega)$$

$$+ c(t)\ \exp\left[\lambda \int_0^t c(\tau)b(\tau,\omega)d\tau\right] \int_0^t \exp\left[-\lambda \int_0^\tau c(\sigma)b(\sigma,\omega)d\sigma\right] F(\tau,\omega)b(\tau,\omega)d\tau. \tag{3.39}$$

Using Equation (3.35), and recalling that $\lambda = -1$, we have

$$y(t,\omega) = F(t,\omega)$$

$$-\exp\left[-\int_0^t \alpha(\tau,\omega)d\tau\right] \int_0^t \exp\left[\left(\int_0^\tau \alpha(\sigma,\omega)d\sigma\right) + s(t-\tau)\right] \alpha(\tau,\omega)F(\tau,\omega)\ d\tau. \tag{3.40}$$

In Equations (3.39) and (3.40), the resolvent kernel of the integral equation (3.34) can be identified.

$$\Gamma(t,\tau,\omega) = c(t)\ \exp\left[\lambda \int_0^t c(\sigma)b(\sigma,\omega)d\sigma\right] \exp\left[-\lambda \int_0^\tau c(\sigma)b(\sigma,\omega)d\sigma\right] b(\tau,\omega), \tag{3.41}$$

or

$$\Gamma(t,\tau,\omega) = \exp\left[-\int_0^t \alpha(\sigma,\omega)\ d\sigma\right] \exp\left[s(t-\tau)\int_0^\tau \alpha(\sigma,\omega)\ d\sigma\right] \alpha(\tau,\omega)\ . \tag{3.42}$$

If $\alpha(\tau,\omega)$ were an ordinary function of time instead of a random function, Equations (3.39) and (3.40) would be the complete solutions to the problem. In the case of stochastic coefficients, we are still faced with the difficult problem of computing the statistical measures of $y(t,\omega)$. We shall investigate the computation of the expected value of $y(t,\omega)$ first. Averaging Equation (3.40), we have

$$< y(t,\omega) > = < F(t,\omega) >$$

$$- < \exp\left[-\int_0^t \alpha(\tau,\omega)d\tau\right] \int_0^t \exp\left[\int_0^t \alpha(\sigma,\omega)d\sigma\right] \alpha(\tau,\omega)> \exp[s(t-\tau)] < F(\tau,\omega)> d\tau .$$

(3.43)

The difficulty arises in computing the following expected value:

$$< \exp\left[-\int_0^t \alpha(\tau,\omega)d\tau\right] \int_0^t \exp\left[\int_0^\tau \alpha(\sigma,\omega)d\sigma\right] \alpha(\tau,\omega) > \quad .$$

(3.44)

This expected value is a nonlinear function of $\alpha(t,\omega)$ and we must resort either to nonlinear transformation techniques or power series expansion of Equation (3.44). To shorten the notation, let

$$\gamma(t,\omega) = \int_0^t \alpha(\tau,\omega)\, d\tau \quad .$$

(3.45)

Then, the power series expansion of (3.43) becomes:

$$< y(t,\omega) > =$$

$$< F(t,\omega) >$$

$$- \int_0^t \exp[s(t-\tau)] < \exp[\gamma(\tau,\omega)-\gamma(t,\omega)]\ \alpha(\tau,\omega) > < F(\tau,\omega) > d\tau$$

$$= < F(t,\omega) >$$

$$- \int_0^t \exp[s(t-\tau)] < \alpha(\tau,\omega)\ [1 + (\gamma(\tau,\omega)-\gamma(\tau,\omega)) + \tfrac{1}{2}\ (\gamma(\tau,\omega)-\gamma(\tau,\omega))^2 +$$

$$\ldots + \frac{1}{n!}\ (\gamma(\tau,\omega)-\gamma(t,\omega))^n > < F(\tau,\omega) > d\tau$$

$$= < F(t,\omega) >$$

$$- \int_0^t \exp[s(t-\tau)] \sum_{n=0}^{\infty} \frac{1}{n!} < \alpha(\tau,\omega)[\gamma(\tau,\omega)-\gamma(t,\omega)]^n > < F(\tau,\omega) > d\tau \quad .$$

(3.46)

Conditions for the convergence of the power series expansion (3.46) can be easily established. By an argument similar to the one used in section 3.2, it can be seen that (3.46) converges even when $\alpha(\tau,\omega)$ is a Gaussian process. The first two terms of the power series expansion are quite easy to compute:

$$< y(t,\omega) > \approx < F(t,\omega) > - \int_o^t \exp\,[s(t-\tau)] < \alpha(\tau,\omega) > < F(\tau,\omega) > d\tau$$

$$- \int_o^t \int_o^\tau \exp\,[s(t-\tau)]\, R_{\alpha\alpha}(\tau,\sigma) < F(\tau,\omega) > d\tau\, d\sigma$$

$$+ \int_o^t \int_o^t \exp\,[s(t-\tau)]\, R_{\alpha\alpha}(\tau,\sigma) < F(\tau,\omega) > d\tau\, d\sigma \ . \tag{3.47}$$

(In using $R_{\alpha\alpha}(\tau,\sigma)$ for $< \alpha(\tau,\omega)\ \alpha(\sigma,\omega) >$ we have implicitly assumed that $\alpha(\tau,\omega)$ is a real process. This assumption is not essential and can be easily removed when $\alpha(\tau,\omega)$ is a complex process.) Of course, if $\alpha(\tau,\omega)$ is a process with zero mean, the integral containing $< \alpha(\tau,\omega) >$ drops out and the approximation becomes:

$$< y(t,\omega) > \approx < F(t,\omega) > - \int_o^t \int_o^\tau \exp\,[s(t-\tau)]\, R_{\alpha\alpha}(\tau,\sigma) < F(\tau,\omega) > d\tau\, d\sigma$$

$$+ \int_o^t \int_o^t \exp\,[s(t-\tau)]\, R_{\alpha\alpha}(\tau,\sigma) < F(\tau,\omega) > d\tau\, d\sigma \ . \tag{3.48}$$

From (3.47) or from (3.48), we can make some observations. First, if $< F(t,\omega) > = 0$, then the mean of the solution $< y(t,\omega) = 0$. On the other hand, if $< F(t,\omega) > \neq 0$, the contribution from the random

parameter $\alpha(t,\omega)$ does not vanish even if $< \alpha(t,\omega) > = 0$. This follows from the fact that

$$\int_0^t \int_0^\tau \exp[s(t-\tau)]\, R(\tau,\sigma) < F(\tau,\omega) > d\tau\, d\sigma \neq$$

$$\int_0^t \int_0^t \exp[s(t-\tau)]\, R(\tau,\sigma) < F(\tau,\omega) > d\tau\, d\sigma \quad . \tag{3.49}$$

Similarly, higher order terms of the power series expansion (3.46) also contribute to the mean value of $< y(t,\omega) >$.

Next, we assume that $F(t,\omega)$ has zero mean. Then, the auto-covariance of $y(t,\omega)$ is obtained by using (3.40) and (3.45):

$$R_{yy}(t_1,t_2) = < y(t_1,\omega)\, y^*(t_2,\omega) >$$

$$= R_{FF}(t_1,t_2)$$

$$- \int_0^{t_1} \exp[s(t_1-\tau)] < \exp[\gamma(\tau,\omega)-\gamma(t_1,\omega)]\alpha(\tau,\omega) > R_{FF}(\tau,t_2)d\tau$$

$$- \int_0^{t_2} \exp[s(t_2-\tau)]^* < \exp[\gamma(\tau,\omega)-\gamma(t_2,\omega)]^* \alpha^*(\tau,\omega) > R_{FF}(t_1,\tau)d\tau$$

$$+ \int_0^{t_1}\int_0^{t_2} \exp[s(t_1-\tau)-s^*(t_2-\sigma)]$$

$$< \exp\{[\gamma(\tau,\omega)-\gamma(t_1,\omega)] + [\gamma(\sigma,\omega)-\gamma(t_2,\omega)]^*\}\ \alpha(\tau,\omega)\ \alpha^*(\sigma,\omega) >$$

$$R_{FF}(\tau,\sigma)\ d\tau\ d\sigma \quad .$$

(3.50)

The two center terms can be expanded in the same manner as in (3.46). An analogous expansion of the last term can also be made.

$$< \exp\{[\gamma(\tau,\omega)-\gamma(t_1,\omega)] + [\gamma(\sigma,\omega)-\gamma(t_2,\omega)]^*\}\alpha(\tau,\omega)\ \alpha^*(\sigma,\omega) >$$

$$= < \alpha(\tau,\omega)\alpha^*(\sigma,\omega)\{1 + [\gamma(\tau,\omega)-\gamma(t_1,\omega)] + [\gamma(\sigma,\omega) - \gamma(t_2,\omega)]^*$$

$$+ \tfrac{1}{2}\,[\gamma(\tau,\omega) - \gamma(t_1,\omega) + \gamma^*(\sigma,\omega) - \gamma^*(t_2,\omega)]^2 \ . \ . \ .$$

$$+ \frac{1}{n!}\,[\gamma(\tau,\omega) - \gamma(t_1,\omega) + \gamma^*(\sigma,\omega) - \gamma^*(t_2,\omega)]^n \ . \ . \ . \ >$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} < \alpha(\tau,\omega)\alpha^*(\sigma,\omega)[\gamma(\tau,\omega) - \gamma(t_1,\omega) + \gamma^*(\sigma,\omega) - \gamma^*(t_2,\omega)]^n >$$

(3.51)

The expansion involves higher and higher moments of $\alpha(t,\omega)$ but does not involve repeated integration as the Neumann series expansion does. The power series expansion (3.51) demonstrates that knowledge of all the moments of stochastic coefficient $\alpha(t,\omega)$ is required for computation of the correlation function $R_{yy}(t_1,t_2)$. We must compute the moments

$$< \alpha(\tau,\omega)\, \alpha^*(\sigma,\omega)[\gamma(\tau,\omega) - \gamma(t_1,\omega) + \gamma^*(\sigma,\omega) - \gamma^*(t_2,\omega)]^n > ,$$

where γ's are given by (3.45). Obviously,

$$< \alpha(\tau,\omega)\, \alpha^*(\sigma,\omega)[\gamma(\tau,\omega) - \gamma(t_1,\omega) + \gamma^*(\sigma,\omega) - \gamma^*(t_2,\omega)]^n >$$

$$\neq < \alpha(\tau,\omega)\, \alpha^*(\sigma,\omega) > < [\gamma(\tau,\omega) - \gamma(t_1,\omega) + \gamma^*(\sigma,\omega) - \gamma^*(t_2,\omega)]^n >.$$

(3.52)

If $\alpha(t,\omega)$ is a Gaussian process, then the higher moments of the process can be computed from the second-order statistics. Hence, a complete solution can be obtained from the knowledge of the second-order statistics.

If $\alpha(t,\omega)$ is a process with zero mean, the simplest approximation of $R_{yy}(t_1,t_2)$, which still takes into account the stochastic coefficient, is obtained by disregarding all but the first terms in the power series expansion of (3.50)

$$R_{yy}(t_1,t_2) \approx R_{FF}(t_1,t_2) + \int_0^{t_1}\int_0^{t_2} e^{-s(t_1-\tau)-s^*(t_2-\tau)}\, R_{\alpha\alpha}(\tau,\sigma) R_{FF}(\tau,\sigma)\, d\tau d\sigma$$

(3.53)

3.5 n-Term Degenerate Kernel*

In this section, the integral equation

$$y(t,\omega) = F(t,\omega) + \lambda \int_0^t K(t,\tau,\omega)\, y(\tau,\omega)\, d\tau \tag{3.54a}$$

with the kernel

$$K(t,\tau,\omega) = \sum_{k=1}^{n} c_k(t)\, b_k(\tau,\omega) \tag{3.54b}$$

is investigated. We assume a solution of the form

$$y(t,\omega) = F(t,\omega) + \sum_{i=1}^{n} A_i(t,\omega)\, c_i(t) \tag{3.54c}$$

Substituting (3.54c) into (3.54a), we get

$$\sum_{i=1}^{n} A_i\, c_i(t) = \lambda \int_0^t \sum_{k=1}^{n} \left[\, c_k(t)\, b_k(\tau,\omega) \,\right] F(\tau,\omega)\, d\tau$$

$$+ \lambda \int_0^t \left[\sum_{k=1}^{n} c_k(t)\, b_k(\tau,\omega) \right] \sum_{j=1}^{n} A_j\, c_j(\tau)\, d\tau \quad . \tag{3.55}$$

---

*Chapter II, Kantorovich and Krylov (1958)

By equating the coefficients of $c_i(t)$, we obtain n simultaneous integral equations for the unknowns A_i:

$$A_i - \lambda \sum_{j=1}^{n} \int_0^t A_j \, b_i(\tau,\omega) \, c_j(\tau) \, d\tau$$

$$= \lambda \int_0^t b_i(\tau,\omega) \, F(\tau,\omega) \, d\tau \qquad (i = 1,2,.....n) \quad . \qquad (3.56)$$

To simplify algebra let:

$$b_i(\tau,\omega) \, c_j(\tau) = B_{ij} \qquad (3.57a)$$

and

$$\int_0^t b_i(\tau,\omega) \, F(\tau,\omega) \, d\tau = f_i \quad . \qquad (3.57b)$$

With this notation, the simultaneoue equations (3.56) can be written as:

$$A_i - \lambda \int_0^t \sum_{j=1}^{n} (A_j B_{ij}) \, d\tau = \lambda f_i \qquad (i = 1,2,....n) \quad ,$$

(3.58)

which can be expressed in the matrix form as:

$$\underline{A} - \lambda \int_0^t \underline{B}\,\underline{A}\, d\tau = \lambda \underline{f} \quad . \tag{3.59}$$

To find the unknown functions A_i, we must solve the matrix integral Equation (3.59). Because the integral Equation (3.54a) is a special case of the more general matrix formulation of the problem, the discussion of further details of the solution are postponed until the next section.

3.6 State Function Formulation

State function formulation has been widely used in modern control system theory and in formulation of Lagrange's equations in classical mechanics. For a system that may be represented by linear differential equations, the state space equations are

$$\dot{\underline{y}}(t,\omega) = \underline{a}(t,\omega)\,\underline{y}(t,\omega) + \underline{U}\,\underline{u}(t,\omega) \quad , \tag{3.60}$$

where $\underline{y}(t,\omega)$ is an n-dimensional state vector, $\underline{a}(t,\omega)$ an n x n system matrix, $\underline{U}$ is an n x r control matrix and $\underline{u}(t,\omega)$ is an r-dimensional control vector. Here, the terminology of control system theory has been used. General mathematical terminology or terminology from classical mechanics could have been used equally well.

Let each member of the system matrix be separable in a deterministic term and a random term. In particular, let the members of the system matrix be of the form

$$a_{ij}(t,\omega) = \beta_{ij} + \alpha_{ij}(t,\omega) \quad , \tag{3.61}$$

where β_{ij} is a constant and the $\alpha_{ij}(t,\omega)$ are stochastic processes. Because β_{ij} is assumed to be a constant, the deterministic part of (3.60) is a linear time-invariant differential equation. The solution to the deterministic differential equation can be expressed in terms of the state transition matrix $\underline{\Phi(t)}$. The methods of construction of the state transition matrices (STM's) for linear time-invariant systems are discussed in several texts on control system theory (Schultz and Melsa, 1967; Zadeh and Desoer, 1963; and De Russo, Roy and Close, 1965). Therefore, it is assumed that the state transition matrix for the deterministic part of the state space equation (3.60) is known. Using the state transition matrices, the state space equations can be converted into a matrix integral equation

$$\underline{y(t,\omega)} = \underline{F(t,\omega)} + \lambda \int_0^t \underline{K(t,\tau,\omega)}\ \underline{y(\tau,\omega)}\ d\tau \quad , \tag{3.62}$$

where $\underline{y(t,\omega)}$ and $\underline{F(t,\omega)}$ are n-dimensional vectors and $\underline{K(t,\tau,\omega)}$ is an n x n matrix. $\underline{F(t,\omega)}$ is the solution of the deterministics part (part with deterministic coefficients) of the state space equations. The kernel of the integral equation is

$$\underline{K(t,\tau,\omega)} = \underline{\Phi(t-\tau)}\ \underline{\alpha(\tau,\omega)} \quad . \tag{3.63}$$

All the matrices in (3.73) are n x n matrices. $\underline{\Phi(t-\tau)}$ is the state transition matrix. For linear time-invariant differential equations, the state transition matrix $\underline{\Phi(t)}$ has the following useful properties (Schultz and Melsa, 1967):

1) It is nonsingular for all finite values of t.

2) $\underline{\Phi(t_1)}\ \underline{\Phi(t_2)} = \underline{\Phi(t_1 + t_2)}$

3) $\underline{\Phi(t)^{-1}} = \underline{\Phi(-t)}$

4) $\underline{\Phi(t)^{n}} = \underline{\Phi(nt)}$. (3.64)

Using the second property, the kernel of the integral equation becomes

$$K(t,\tau,\omega) = \underline{\Phi(t-\tau)}\ \underline{\alpha(\tau,\omega)}$$

$$= \underline{\Phi(t)}\ \underline{\Phi(-\tau)}\ \underline{\alpha(\tau,\omega)}$$

$$= \underline{\Phi(t)}\ \underline{b(\tau,\omega)} \quad , \tag{3.65}$$

Here, for simplicity, the product $\underline{\Phi(-\tau)}\ \underline{\alpha(\tau,\omega)}$ is denoted by $\underline{b(\tau,\omega)}$. Now, the integral equation (3.62) has a degenerate kernel and can be solved by assuming a solution of the following form:

$$\underline{y(t,\omega)} = \underline{F(t,\omega)} + \underline{\Phi(t)}\ \underline{A(t,\omega)} \quad , \tag{3.66}$$

where $\underline{A(t,\omega)}$ is an n-dimensional vector. Substituting Equation (3.66) into Equation (3.62), we have:

$$\underline{\Phi(t)}\ \ \underline{A(t,\omega)} = \lambda \int_0^t \underline{\Phi(t)}\ \ \underline{b(\tau,\omega)}\ \ \underline{F(\tau,\omega)}\ \ d\tau$$

$$+ \lambda \int_0^t \underline{\Phi(t)}\ \ \underline{b(\tau,\omega)}\ \ \underline{\Phi(\tau)}\ \underline{A(t,\omega)}\ d\tau. \qquad (3.67)$$

Premultiplying both sides of Equation (3.67) by $\Phi^{-1}(t)$ [$\Phi^{-1}(t)$ exists according tc the first property of Equation (3.64)] and rearranging gives

$$\left[\underline{A(t,\omega)} - \lambda \left\{ \int_0^t \underline{b(\tau,\omega)}\ \underline{\Phi(\tau)}\ \underline{A(\tau,\omega)}\ d\tau \right\} \right] = \lambda \int_0^t \underline{b(\tau,\omega)}\ \underline{F(\tau,\omega)}\ d\tau. \qquad (3.68)$$

This matrix integral equation must be solved for $\underline{A(t,\omega)}$. This can be done by again using Neumann series expansion; i.e.,

$$\underline{A(t,\omega)} = \sum_{n=0}^{\infty} \underline{A_n(t,\omega)}\ , \qquad (3.69)$$

where the $A_n(t,\omega)$ terms are given by the following iterative solution:

$$A_o(t,\omega) = \int_0^t \underline{b(\tau,\omega)}\ \ \underline{F(\tau,\omega)}\ \ d\tau\ , \qquad (3.70a)$$

and

$$A_1(t,\omega) = \int_o^t \underline{b(\tau,\omega)}\ \underline{\phi(\tau)}\ \underline{A_o(\tau,\omega)}\ d\tau$$

$$+ \int_o^t \int_o^\tau \underline{b(\tau,\omega)}\ \underline{\phi(\tau)}\ \underline{b(\sigma,\omega)}\ \underline{F(\sigma,\omega)}\ d\tau\ d\sigma\ , \qquad (3.70b)$$

or, in general,

$$\underline{A_n(t,\omega)} = \int_o^t \underline{b(\tau,\omega)}\ \underline{\phi(\tau)}\ \underline{A_{n-1}(\tau,\omega)}\ d\tau. \qquad (3.70c)$$

Sufficient conditions for the convergence of the Neumann series (3.69) are given by Theorem 1 of section 3.2.

Using (3.66), the expected value of $\underline{y(t,\omega)}$ is:

$$< y(t,\omega) > = < F(t,\omega) > + \underline{\phi(t)}\ < A(t,\omega) >$$

$$= < F(t,\omega) > + \underline{\phi(t)} \sum_{n=0}^{\infty} < \underline{A_n(t,\omega)} > ,$$

where $\underline{A_n(t,\omega)}$ is given by iterative integrals (3.70a), (3.70b) and (3.70c). Here again, it can be observed that if $< F(t,\omega) > \neq 0$, $< \underline{y(t,\omega)} >$ does not necessarily vanish when $< \underline{\alpha(t,\omega)} > = 0$ because expressions for $< \underline{A_n(t,\omega)} >$ contain higher moments of $\underline{\alpha(t,\omega)}$ which may not vanish. The first two terms in the Neumann series, $A_o(t,\omega)$ and $A_1(t,\omega)$, can be computed from the second-order statistics; for

higher order terms, we need more complete statistical knowledge of the stochastic parameter matrix $\underline{\alpha(t,\omega)}$. For the approximation which uses the second-order statistics of the stochastic parameters, we have:

$$\langle \underline{y(t,\omega)} \rangle \approx \langle \underline{F(t,\omega)} \rangle + \int_0^t \underline{\Phi(t-\tau)} \langle \underline{\alpha(\tau,\omega)} \rangle \langle \underline{F(\tau,\omega)} \rangle d\tau$$

$$+ \int_0^t \int_0^\tau \underline{\Phi(t-\tau)} \langle \underline{\alpha(\tau,\omega)\ \Phi(\tau-\sigma)\ \alpha(\sigma,\omega)} \rangle \langle F(\sigma,\omega) \rangle d\tau d\sigma. \tag{3.71a}$$

If the coefficient matrix $\alpha(\tau,\omega)$ has zero mean; i.e.,

$$\langle \underline{\alpha(\tau,\omega)} \rangle = 0 ,$$

then, (3.71a) simplifies further and it becomes:

$$\langle y(t,\omega) \rangle \approx \langle \underline{F(t,\omega)} \rangle$$

$$+ \int_0^t \int_0^\tau \Phi(t-\tau) \langle \alpha(\tau,\omega)\ \Phi(\tau-\sigma)\ \alpha(\sigma,\omega) \rangle \langle F(\sigma,\omega) \rangle d\tau d\sigma. \tag{3.71b}$$

The Equation (3.71b) demonstrates how the mean value of $y(t,\omega)$ depends on the second moments of the coefficient matrix. Of course, in this approximation, we neglected the functions which contain higher moments of the coefficient matrix.

The expression for the covariance matrix is obtained by post-multiplying $\underline{y}(t_1,\omega)$ by the complex conjugate transpose of $\underline{y}(t_2,\omega)$ (hermetian conjugate denoted by $\dagger$) and averaging; i.e.;

$$\underline{R}_{yy}(t_1,t_2) = \langle \underline{y}(t_1,\omega)\, \underline{y}^{\dagger}(t_2,\omega) \rangle$$

$$= \langle (\underline{F}(t_1,\omega)) + \underline{\Phi}(t_1)\, \underline{A}(t_1,\omega))(\underline{F}(t_2,\omega) + \underline{\Phi}(t_2)\, \underline{A}(t_2,\omega))^{\dagger} \rangle$$

$$= \langle \underline{F}(t_1,\omega)\, \underline{F}^{\dagger}(t_2,\omega) \rangle + \langle \underline{\Phi}(t_1)\, \underline{A}(t_1,\omega)\, \underline{F}^{\dagger}(t_2,\omega) \rangle$$

$$+ \langle \underline{F}(t_1,\omega)\, \underline{A}^{\dagger}(t_2,\omega)\, \underline{\Phi}^{\dagger}(t_2) \rangle$$

$$+ \langle \underline{\Phi}(t_1)\, \underline{A}(t_1,\omega)\, \underline{A}^{\dagger}(t_2,\omega)\, \underline{\Phi}^{\dagger}(t_2) \rangle$$

$$= \underline{R}_{FF}(t_1,t_2) + \underline{\Phi}(t_1) \langle \underline{A}(t_1,\omega)\, \underline{F}^{\dagger}(t_2,\omega) \rangle$$

$$+ \langle \underline{F}(t_1,\omega)\, \underline{A}^{\dagger}(t_2,\omega) \rangle \underline{\Phi}^{\dagger}(t_2)$$

$$+ \underline{\Phi}(t_1) \langle \underline{A}(t_1,\omega)\, \underline{A}^{\dagger}(t_2,\omega) \rangle \underline{\Phi}^{\dagger}(t_2) \quad . \tag{3.72}$$

If only the first terms of the power series expansions are used, an approximate expression for the covariance matrix $\underline{R_{yy}(t_1,t_2)}$ is

$$\underline{R_{yy}(t_1,t_2)} \approx \underline{R_{FF}(t_1,t_2)}$$

$$+ \underline{\Phi(t_1)} \int_0^t \underline{\Phi^{-1}(\tau)} \, \underline{< \alpha(\tau,\omega) >} \, \underline{R_{FF}(\tau_1,t_2)} \, d\tau$$

$$+ \left(\int_0^{t_2} \underline{R_{FF}(t_1,\tau)} \, \underline{< \alpha^\dagger(\tau,\omega) >} \, \underline{\Phi^{-1^\dagger}(\tau)} \, d\tau \right) \underline{\Phi^\dagger(t_2)}$$

$$+ \underline{\Phi(t_1)} \left(\int_0^{t_1} \int_0^{t_2} \underline{\Phi^{-1}(\tau)} \, \underline{R_{\alpha^2F^2}(\tau,\sigma)} \, \underline{\Phi^{-1^\dagger}(\sigma)} \, d\tau \, d\sigma \right) \underline{\Phi^\dagger(t_2)} \, , \tag{3.73a}$$

where

$$R_{\alpha^2F^2}(\tau,\sigma) = < [\alpha(\tau,\omega) \, F(\tau,\omega)] \, [\alpha(\sigma,\omega) \, F(\sigma,\omega)]^\dagger > \, . \tag{3.73b}$$

The approximate expression for $\underline{R_{yy}(t_1,t_2)}$ simplifies further if

$$< \alpha(\tau,\omega) > \; = \; < \alpha^\dagger(\tau,\omega) > \; = \; 0 \quad .$$

Then, (3.73a) becomes:

$$R_{yy}(t_1,t_2) \approx R_{FF}(t_1,t_2) \; +$$

$$+ \underline{\Phi(t_1)} \left(\int_0^{t_1} \int_0^{t_2} \underline{\Phi^{-1}(\tau)} \, \underline{R_{\alpha^2F^2}(\tau,\sigma)} \, \underline{\Phi^{-1^\dagger}(\sigma)} \, d\tau \, d\sigma \right) \underline{\Phi^\dagger(t_2)} \, . \tag{3.74}$$

The equation (3.74) also shows some interesting results. The first term in (3.74) is the contribution of the deterministic part of the differential equation. The next term is the simplest term which takes into account the randomness of the coefficients. To compute additional terms in the expansion, we need to know the higher moments of the coefficients. If only the second-order statistics of the coefficients are known, (3.74) is as far as we can go without additional knowledge.

In the state space formulation, one has considerable freedom in the selection of the coefficient matrix representation. The evaluation of the truncation error is greatly simplified if one can select a representation for the coefficient matrix which will make the Neumann series expansion of $A(t,\omega)$ an alternating series. Then, for a convergent series, the truncation error is smaller than the first term neglected, provided the norm of each term is smaller than the norm of its preceding term.

3.7 General Remarks

In using the degenerate kernel approach for solving the Volterra integral equation, we still have to resort to Neumann series expansion. From the standpoint of computational difficulty, there is no essential difference between the straight Neumann series expansion and the degenerate kernel approach. Selection of the method depends on the physical problem and computational convenience. For example, in the study of wave propagation in random media, iterative integrals of the Neumann series expansion have the convenient interpretation of repeated scattering of waves. On the other hand, in the case of a control system problem, the deterministic transition matrix may be

already known and, therefore, the degenerate kernel method may be convenient. Both methods of solution converge under the same general conditions. Hence, we have two complementary methods for solving the stochastic Volterra integral equations. Other methods, such as the Fredholm method and the Hilbert-Schmidt method, do not offer any computational advantages. To use the Hilbert-Schmidt method, one has to solve a pair of first kind integral equations. This problem is no easier than solving an integral equation of the second kind. Fredholm theory has been very important in the development of the classical integral equation theory, but to use the Fredholm method for construction of the resolvent kernel is prohibitively difficult in practice.

Many aribtrary kernels can be approximated by the degenerate kernels. A Taylor series expansion of an arbitrary kernel can be used for approximating it with a degenerate kernel. For example, a kernel $\sin(t\,\tau)$ can be approximated by

$$\sin(t\,\tau) \approx t\,\tau - \frac{t^3\tau^3}{3!} + \frac{t^5\tau^5}{5!} \quad . \tag{3.75}$$

A Fourier series expansion of the arbitrary kernel or special interpolation devices can also be used. The use of the method of moments for solving integral equations is equivalent to the replacement of an arbitrary kernel by a degenerate kernel. In the case of deterministic integral equations, the estimates of errors caused by the replacement of a given kernel by a degenerate one are known. One such theorem, slightly modified for stochastic application, is stated below (Kantorovich and Krylov, 1958).

Theorem 2 . Let there be two kernels, $k(t,\tau,\omega)$ and $K(t,\tau,\omega)$ and let it be known that

$$\int_{t_o}^{t_1} |K(t,\tau,\omega) - k(t,\tau,\omega)| \, d\tau < h \tag{3.76}$$

almost everywhere P (a.e.P.), and that the resolvent kernel $\gamma(t,\tau,\omega,\lambda)$ of the equation with kernel $k(t,\tau,\omega)$ satisfies the inequality a.e.P.

$$\int_{t_o}^{t_1} |\gamma(t,\tau,\omega,\lambda)| \, d\tau < B \quad , \tag{3.77}$$

and also that

$$|f(t) - f_1(t)| < \eta \text{ a.e.P.} \quad . \tag{3.78}$$

Then, if the following condition is satisfied a.e.P.:

$$1 - |\lambda| \, h(1 + |\lambda| \, B) > 0 \quad , \tag{3.79}$$

the equation

$$y(t,\omega) - \lambda \int_{t_o}^{t_1} K(t,\tau,\omega) \, y(\tau,\omega) \, d\tau = f(t,\omega) \tag{3.80}$$

has a unique solution $y(t,\omega)$ a.e.P., and the difference between this solution and the solution $\hat{y}(t,\omega)$ of the equation

$$\hat{y}(t,\omega) - \lambda \int_{t_o}^{t_1} k(t,\tau,\omega)\, y(\tau,\omega)\, d\tau = f_1(\tau,\omega) \qquad (3.81)$$

is smaller a.e.P. than

$$|y(t,\omega) - y(t,\omega)| < \frac{N|\lambda|\, h(1 + |\lambda|\, B)^2}{1 - |\lambda|\, h\, (1 + |\lambda|\, B} + \eta\, (1 + |\lambda|\, B), \qquad (3.82)$$

where N is the upper bound of $|f(t,\omega)|$ a.e.P.

Proof is a trivial modification of the deterministic proof given by Kantorovich and Krylov (1958). The theorem is less than satisfactory for many interesting processes; however, the hypotheses of the theorem are too restrictive and the conclusion of the theorem should give the mean square approximation error. To make the above theorem applicable to the Volterra integral, we define the Volterra kernels as follows:

$$K_v(t,\tau,\omega) = K(t,\tau,\omega) \qquad \tau < t$$
$$= 0 \qquad \tau \geq t \qquad (3.83)$$

3.8 Conclusions

This chapter presents two methods for obtaining expressions for the ensemble average and covariance of the solution of an n^{th} order differential equation with stochastic coefficients. In both cases, it is assumed that the coefficients of the differential equation are

separable into deterministic and stochastic coefficients. The problem now becomes a problem of solving a Volterra integral equation with a stochastic kernel. In the first method, the problem is solved by the Neumann series expansion. The Neumann series expansion is a series of iterated integrals. A previously computed term is used to compute higher order terms. Sufficient conditions for the convergence of the Neumann series are given. The Neumann series is used to find the resolvent kernel of the stochastic integral equation. The ensemble averages and covariance functions of the solution are expressed in terms of the resolvent kernel and the corresponding statistical measures of the input process. The kernels of these integral expressions for the statistical measures of the solution can be interpreted as the stochastic Green's functions.

In the second method, it is assumed that the deterministic part of the differential equation is time invariant. Then, in many cases, the kernels of the integral equation are degenerate kernels. In these cases, we have a slightly different method for solving the Volterra integral equation. We still have to resort to Neumann series expansion for the complete solution of the integral equation. From the standpoint of computational difficulty and convergence of the solution, there is no essential difference between the two methods.

However, we have two complementary methods of solving the problem. Applying these methods to some very simple equations and just computing the first couple terms of the expansion, we are able to observe some interesting results. Even if the stochastic coefficients have zero mean, their contribution to the mean value of the solution does not

vanish. This demonstrates that the average of the solution of a differential equation with stochastic coefficients is not necessarily the same as the solution of the averaged equation. Both methods of solution also show that expressions for the second-order statistics of the solution of linear stochastic differential equations requires knowledge of all the moments of stochastic coefficients. An exceptional case is that in which the stochastic coefficients are Gaussian processes. Then, the knowledge of the second-order statistics is sufficient for the complete solution.

In both cases, the computation is simplified if the state space formulation is used. The use of the state space formulation has the further advantage that it connects modern control system theory with this work.

The concepts developed in this chapter will be generalized in the next chapter to partial differential equations and applied to the propagation of the scalar wave function in a randomly time-varying medium.

CHAPTER IV

WAVE PROPAGATION IN A RANDOMLY TIME AND SPACE VARYING MEDIUM

4.1 Introduction

The problem of propagation of an electromagnetic wave in a random continuous media has been studied extensively by many workers;* however, in almost all cases, attention has been limited to a random medium with space-varying statistical properties. In these studies, a so-called "quasimonochromatic" solution has been assumed.** This quasimonochromatic assumption essentially neglects the time-varying properties of the medium. In many cases, this assumption may very well be correct, for example, in the case of the propagation of light through frosted glass. In cases when one deals with wave propagation through hot or very energetic media such as the atmospheres of stars or plasmas, the quasimonochromatic assumption is clearly incorrect. In other cases, such as in the study of synchronization of spatially separated frequency standards, doppler broadening of radar signals or interaction of two signals in a nonlinear medium, the small frequency shifts caused by the randomly time-varying medium may be the important questions under study. In all cases, the validity of the quasimonochromaticity assumption should be verified.

---

*See the references in section 1.3.

**One assumes that the solution of the wave equation is essentially a sinusoid at a single frequency, if the source is a sinusoid at a single frequency.

It is assumed that the dielectric permittivity $\epsilon(t,P,\omega)$ is a random function of time t and position P. It is further assumed that it can be separated into a constant term ϵ_o plus a randomly space- and time-varying term. The wave equation (4.1) for the electric field $\bar{E}$ can be derived from the pair of Maxwell's equations that connect the electric and magnetic fields*

$$\nabla \times \nabla \times \bar{E} + \frac{\partial^2}{\partial t^2} (\mu_o \epsilon \bar{E}) = - \mu_o \frac{\partial \bar{J}}{\partial t} , \qquad (4.1)$$

where $\bar{J}$ is the current density.

In the quasimonochromatic assumption, it is assumed that

$$\frac{\partial^2}{\partial t^2} \mu_o \epsilon \bar{E} \approx \mu_o \epsilon \frac{\partial^2}{\partial t^2} \bar{E} , \qquad (4.2a)$$

and that the equation (4.1) can be written, in the current free region,

$$\nabla \times \nabla \times \bar{E}_o - s^2 \mu_o \epsilon \bar{E}_o = 0 , \qquad (4.2b)$$

where $\bar{E} = \bar{E}_o e^{ist}$.

These assumptions are avoided in this dissertation. Instead, it is assumed that the spacial gradient of the dielectric permittivity over the distance of one wavelength is small:

---

*An overbar - denotes vectors.

CHAPTER IV

WAVE PROPAGATION IN A RANDOMLY TIME AND SPACE VARYING MEDIUM

4.1 Introduction

The problem of propagation of an electromagnetic wave in a random continuous media has been studied extensively by many workers;* however, in almost all cases, attention has been limited to a random medium with space-varying statistical properties. In these studies, a so-called "quasimonochromatic" solution has been assumed.** This quasimonochromatic assumption essentially neglects the time-varying properties of the medium. In many cases, this assumption may very well be correct, for example, in the case of the propagation of light through frosted glass. In cases when one deals with wave propagation through hot or very energetic media such as the atmospheres of stars or plasmas, the quasimonochromatic assumption is clearly incorrect. In other cases, such as in the study of synchronization of spatially separated frequency standards, doppler broadening of radar signals or interaction of two signals in a nonlinear medium, the small frequency shifts caused by the randomly time-varying medium may be the important questions under study. In all cases, the validity of the quasimonochromaticity assumption should be verified.

---

*See the references in section 1.3.

**One assumes that the solution of the wave equation is essentially a sinusoid at a single frequency, if the source is a sinusoid at a single frequency.

It is assumed that the dielectric permittivity $\epsilon(t,P,\omega)$ is a random function of time t and position P. It is further assumed that it can be separated into a constant term ϵ_o plus a randomly space- and time-varying term. The wave equation (4.1) for the electric field $\bar{E}$ can be derived from the pair of Maxwell's equations that connect the electric and magnetic fields*

$$\nabla \times \nabla \times \bar{E} + \frac{\partial^2}{\partial t^2} (\mu_o \epsilon \bar{E}) = - \mu_o \frac{\partial \bar{J}}{\partial t} , \qquad (4.1)$$

where $\bar{J}$ is the current density.

In the quasimonochromatic assumption, it is assumed that

$$\frac{\partial^2}{\partial t^2} \mu_o \epsilon \bar{E} \approx \mu_o \epsilon \frac{\partial^2}{\partial t^2} \bar{E} , \qquad (4.2a)$$

and that the equation (4.1) can be written, in the current free region,

$$\nabla \times \nabla \times \bar{E}_o - s^2 \mu_o \epsilon \bar{E}_o = 0 , \qquad (4.2b)$$

where $\bar{E} = \bar{E}_o e^{ist}$.

These assumptions are avoided in this dissertation. Instead, it is assumed that the spacial gradient of the dielectric permittivity over the distance of one wavelength is small:

---

*An overbar - denotes vectors.

$$|\nabla\epsilon|\ \lambda/\epsilon \ll 1 \quad . \tag{4.3}$$

Then, (4.1) becomes the following wave equation

$$\nabla^2 \bar{E} - \frac{\partial^2}{\partial t^2}\ \mu_o \epsilon \bar{E} = \mu_o \frac{\partial \bar{J}}{\partial t} \quad . \tag{4.4}$$

To keep the basic problem uncluttered of nonessential mathematical complexity, we restrict the analysis to the solution of the following scalar wave equation:

$$\nabla^2 y(t,\bar{r},\omega) - \frac{\partial^2}{\partial t^2}\left(\frac{1}{c^2} + \alpha(t,\bar{r},\omega)\right) y(t,\bar{r},\omega) = x(t,\bar{r},\omega) \tag{4.5}$$

$t\epsilon T$, $\bar{r}\epsilon R^3$ and $\omega\epsilon\Omega$ on the probability space $(\Omega,\mathcal{X},P)$. $\alpha(t,\bar{r},\omega)$ is a stochastic coefficient; that is, a random function of both time and space. $x(t,\bar{r},\omega)$ is the stocnastic source term. The source term could be a deterministic function but, for greater generality, it is assumed to be stochastic. It is assumed that the source term $x(t,\bar{r},\omega)$ and the coefficient $\alpha(t,\bar{r},\omega)$ are statistically independent, For simplicity, it is assumed that $\alpha(t,\bar{r},\omega)$ is a real stochastic process. In order to be able to use some of the results from the theory of wide-sense stationary stochastic processes, it is assumed that both $\alpha(t,\bar{r},\omega)$ and $x(t,\bar{r},\omega)$ are wide-sense stationary or reducible to wide-sense stationary stochastic processes.* This assumption

---

*Discussed in section 2.4.

simplifies the development in that the spectral representations of $\alpha(t,\bar{r},\omega)$ and $x(t,\bar{r},\omega)$ have some useful properties. Wide-sense stationarity only in time t is required. It is not required that $\alpha(t,\bar{r},\omega)$ and $x(t,\bar{r},\omega)$ are wide-sense stationary in their spacial variable $\bar{r}$. The statistical properties of $\alpha(t,\bar{r},\omega)$ and $x(t,\bar{r},\omega)$ may change in an arbitrary manner in the spacial variable as long as $\alpha(t,\bar{r},\omega)$ and $x(t,\bar{r},\omega)$ remain second-order processes and (4.3) remains satisfied. By placing no overly restrictive conditions on the spacial statistics of the medium, physical reality is maintained. The statistical properties of the medium may vary in space in a manner that accommodates the usual physical problems. The randomness of the medium may be restricted to be in a given volume, shown by Figure 4.1, or the mean square amplitude of the fluctuation of $\alpha(t,\bar{r},\omega)$ may vary with height as it is common in the tropospheric communication problems. The reducibility to wide-sense stationarity also allows us to consider physically reasonable problems. The sample functions of $\alpha(t,\bar{r},\omega)$ and $x(t,\bar{r},\omega)$ fluctuate rapidly in time, but their averaged properties vary slowly with time. It is rapid time fluctuation of the sample function $\alpha(t,\bar{r},\omega)$ which makes the quasimonochromatic assumption dubious. The concept of reducibility to stationarity permits slow variation of the statistical properties of $\alpha(t,\bar{r},\omega)$ and $x(t,\bar{r},\omega)$. These slow variations of the statistical properties must be sufficiently slow so that they do not mask the spectral spreading which is caused by the rapid time variation of the sample function $\alpha(t,\bar{r},\omega)$. This condition is satisfied if the power spectrum of the function which modulates the statistical properties of $\alpha(t,\bar{r},\omega)$ and $x(t,\bar{r},\omega)$ contains only frequencies which

are very much lower than the significant frequencies of these stochastic processes. In future work, it would be desirable to remove even the wide-sense stationarity assumption.

The use of a scalar wave equation also limits the generality of the analysis. It neglects the change in polarization due to a random medium, and it is strictly applicable only to the forward scattering of electromagnetic waves. Of course, in case of sound propagation, there is no loss of generality when a scalar wave equation is used.

The statistical measures of interest in this chapter are the power spectral density and coherence functions of the scalar wave function $y(t,\bar{r},\omega)$. The expressions for the spectral density will reveal the spectral spreading caused by a randomly time-varying medium. Coherence functions have been found useful in statistical optics (Born and Wolf, 1964; Beran and Parrent, 1964; O'Neill, 1963; and Mandel and Wolf, 1965) and more recently in general electromagnetic theory (Special Issue on Partial Coherence, 1967). To solve our problem, the differential equation (4.5) is converted into a differential equation for the spectral representations of $y(t,\bar{r},\omega)$, then the methods of Chapter III are applied. Both the Neumann series expansion and degenerate kernel approximations are investigated.

4.2 Spectral Representation of the Scalar Wave Equation

In principle, it is possible to apply the Neumann series solution of section 3.2 directly to the scalar wave equation (4.5) but, after the first few iterations, the time-domain expressions become unmanageable and difficult to interpret. Furthermore, the interesting statistical measure in this case is the power spectrum of the scalar wave

function. Other second-order statistical measures such as variance and mutual coherence functions can be obtained by taking the inverse Fourier transform of the power spectrum. Because there is a simple relation between spectral representation of a random variable and the power spectrum (Chapter II), equation (4.5) is solved for the spectral representation of $y(t,\bar{r},\omega)$. The spectral representation is also called the "integral canonical expansion" of the random functions (Pugachev, 1965).

If the random coefficient $\alpha(t,\bar{r},\omega)$ and the forcing function $x(t,\bar{r},\omega)$ are wide-sense stationary stochastic processes with zero mean, their integral expansions are:

$$\alpha(t,\bar{r},\omega) = \int_{-\infty}^{\infty} A(u,\bar{r},\omega)e^{iut}\,du \tag{4.6a}$$

and

$$x(t,\bar{r},\omega) = \int_{-\infty}^{\infty} X(z,\bar{r},\omega)e^{izt}\,dz\ , \tag{4.6b}$$

where u and z are real variables. The integrals of (4.6) should be written as Stieljes integrals

$$\alpha(t,\bar{r},\omega) = \int_{-\infty}^{\infty} e^{iut}\,d\,A_o(u,\omega) \tag{4.7a}$$

and

$$x(t,\bar{r},\omega) = \int_{-\infty}^{\infty} e^{izt}\,d\,X_o(z,\omega)\ , \tag{4.7b}$$

because d $A_o(u,\omega)$ and d $X_o(z,\omega)$ are stochastic processes with orthogonal increments, and the functions $A_o(u,\omega)$ and $X_o(z,\omega)$ need not be differentiable. For simplicity, the integral expansions (4.6) are used with the understanding that $A(u,\bar{r},\omega)$ and $X(z,\bar{r},\omega)$ may be generalized functions. The assumption that $x(t,\bar{r},\omega)$ and $\alpha(t,\bar{r},\omega)$ have zero mean constitutes no loss of generality, because the mean value of $\alpha(t,\bar{r},\omega)$ can be included with the deterministic part of the coefficient and the non-zero mean value of the source term $x(t,\bar{r},\omega)$ can be easily taken care of by superposition. The wide-sense stationarity assumption constitutes a more serious loss of generality but, without this assumption, the solution of (4.5) becomes very difficult. In future work, it would be desirable to remove this assumption. Actually, the stationarity assumption can be slightly relaxed by assuming that $\alpha(t,\bar{r},\omega)$ are processed reducible to stationary. This will be discussed shortly.

Substituting (4.6a) and (4.6b) into (4.5), then multiplying both sides of the equation by e^{-ist} and integrating, equation (4.5) becomes:

$$\int_{-\infty}^{\infty} dt\, e^{-ist}\, (\nabla^2 y(t,\bar{r},\omega)) \;-\; \frac{1}{c^2} \int_{-\infty}^{\infty} dt\, e^{-ist}\, \frac{\partial^2}{\partial t^2}\, y(t,\bar{r},\omega)$$

$$-\int_{-\infty}^{\infty} dt\, e^{-ist}\, \frac{\partial^2}{\partial t^2} \int_{-\infty}^{\infty} du\, e^{iut}\, A(u,\bar{r},\omega)\; y(t,\bar{r},\omega) \;=$$

$$= \int_{-\infty}^{\infty} dt\, e^{-ist} \int_{-\infty}^{\infty} dz\; X(z,\bar{r},\omega)\, e^{izt}\,, \tag{4.8}$$

where s, u and z are real variables. Interchanging the order of integration and ∇^2 operation, the first term in (4.8) becomes $2\pi\nabla^2 Y(s,\bar{r},\omega)$. Integrating by parts, the second term in (4.8) becomes $2\pi\, s^2/c^2\, Y(s,\bar{r},\omega)$, and the third term becomes

$$s^2 \int_{-\infty}^{\infty} du\, A(u,\bar{r},\omega) \int_{-\infty}^{\infty} dt\, y(t,\bar{r},\omega)\, e^{-i(s-u)t} \tag{4.9a}$$

or

$$2\pi\, s^2 \int_{-\infty}^{\infty} du\, A(u,\bar{r},\omega)\, Y(s-u,\bar{r},\omega) \quad . \tag{4.9b}$$

To obtain these results, the following quantities

$$e^{-ist} \frac{\partial}{\partial t}\, y(t,\bar{r},\omega) \tag{4.10a}$$

$$e^{-ist}\, y(t,\bar{r},\omega) \tag{4.10b}$$

$$e^{-ist} \frac{\partial}{\partial t} \int_{-\infty}^{\infty} du\, A(u,\bar{r},\omega)\, y(t,\bar{r},\omega)\, e^{iut} \tag{4.10c}$$

and

$$e^{-ist} \int_{-\infty}^{\infty} du\, A(u,\bar{r},\omega)\, y(t,\bar{r},\omega)\, e^{iut} \tag{4.10d}$$

must vanish as t approaches $\pm\infty$.

If $x(t,\bar{r},\omega)$ is a wide-sense stationary stochastic process, the terms of (4.10) do not vanish as $t \to \pm \infty$, because wide-sense stationarity of $x(t,\bar{r},\omega)$ implies that the forcing function is active from time $-\infty$ to $+\infty$. This is physically unreasonable. To remove this difficulty, it can be assumed that $x(t,\bar{r},\omega)$ is a process that is "reducible to a stationary process". A stochastic process is said to be <u>reducible to a stationary process</u> if it can be expressed in terms of stationary stochastic processes.* An example of a stochastic process which is reducible to a stationary process is any process of the form

$$x(t,\omega) = g(t)\, z(t,\omega) + f(t) , \tag{4.11}$$

where $z(t,\omega)$ is a stationary stochastic process, and $g(t)$ and $f(t)$ are real nonrandom functions of time. In particular, we may take $f(t)$ to be zero and $g(t)$ can be selected so that the terms of (4.10) vanish as $t \to \pm \infty$. For example, if taken to be

$$g(t) = e^{-kt^2} , \tag{4.12}$$

where k is some small positive number, then $x(t,\bar{r},\omega)$ is a process reducible to a stationary process and terms of (4.10) vanish as $t \to \pm \infty$. (This device is frequently used in the theory of distributions.) The source term becomes

---

*Discussed in section 2.4.

$$\int_{-\infty}^{\infty} dt\, e^{ist} \int_{-\infty}^{\infty} X(z,\bar{r},\omega)\, e^{-zt}\, dz$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} X(z,\bar{r},\omega)\, e^{i(z-s)t}\, dz$$

$$= 2\pi \int_{-\infty}^{\infty} X(s,\bar{r},\omega)\, \delta(z-s)\, dz$$

$$= 2\pi\ X(s,\bar{r},\omega) \quad . \tag{4.13}$$

Collecting the results, the spectral representation of the differential equation (4.5) becomes

$$\nabla^2\, Y(s,\bar{r},\omega) + s^2/c^2\, Y(s,\bar{r},\omega) = X(s,\bar{r},\omega) - s^2 \int_{-\infty}^{\infty} du\, A(u,\bar{r},\omega)\, Y(s-u,\bar{r},\omega) \quad . \tag{4.14}$$

The left-hand side of the equation (4.14) is simply the Helmholtz equation for the spectral representation of the scalar wave function $Y(s,\bar{r},\omega)$. Solutions of the Helmholtz equation are known for a number of different boundary conditions; therefore, for these conditions (4.14) can be expressed as an integral equation:

$$Y(s,\bar{r},\omega) = F(s,\bar{r},\omega) - L^{-1}\, s^2 \int_{-\infty}^{\infty} Y(s-u,\bar{r}',\omega)\, A(u,\bar{r}',\omega)\, du \quad , \tag{4.15}$$

where L^{-1} is the inverse operator for the Helmholtz equation for the appropriate boundary conditions. $F(s,\bar{r},\omega)$ denotes $L^{-1} X(s,\bar{r}',\omega)$; that is, $F(s,\bar{r},\omega)$ is simply the solution of the deterministic Helmholtz equation. The iterative solution is simplified if a change of variable is made in the last integral and the integral equation (4.15) is written as:

$$Y(s,\bar{r},\omega) = F(s,\bar{r},\omega) - L^{-1} s^2 \int_{-\infty}^{\infty} Y(u,\bar{r}',\omega)\, A(s-u,\bar{r}',\omega)\, du \ . \tag{4.16}$$

Now, the scalar wave equation (4.5) has been converted into a Fredholm integral equation for the spectral representation of the scalar wave function. To solve these integral equations, either the Neumann series expansion or the degenerate kernel approximation can be used. For the sake of being more specific and for simplification, attention is restricted to the wave propagation in a spherical coordinate system. The degenerate kernal approximation is considered first.

4.3 Degenerate Kernel Approximation

The integral equation (4.16) becomes an integral equation with a degenerate kernel if the following usual approximations for large R are used:

$$|\bar{r} - \bar{r}'| = R - R' \cos\psi' + 0\left(\frac{1}{R}\right) \tag{4.17a}$$

and

$$\frac{1}{|\bar{r} - \bar{r}'|} = \frac{1}{R} + 0\left(\frac{1}{R^2}\right) \ . \tag{4.17b}$$

With these approximations, (4.16) becomes:

$$Y(s,R,\omega) = F(s,R,\omega)$$

$$+ \int_V \frac{s^2 \exp\left[-i \frac{s}{c} (R - R' \cos \psi')\right]}{4\pi R} \int_{-\infty}^{\infty} Y(u,R',\omega)\, A(s-u,R',\omega)\, du\, dv \,. \tag{4.18}$$

The geometry of the problem is shown in Figure 4.1. V denotes volume integration over the volume occupied by the random medium. This is an integral equation with a degenerate kernel and it can be solved by assuming a solution of the following form:

$$Y(s,R,\omega) = F(s,R,\omega) + B(\omega)\, s^2 \frac{\exp\left[-i \frac{s}{c} R\right]}{4\pi R} \,. \tag{4.19}$$

Substituting (4.19) into (4.18) and solving for $B(\omega)$, we have:

$$B(\omega) = \frac{\int_V \int_{-\infty}^{\infty} \exp\left[i \frac{s}{c} R' \cos \psi'\right] F(u,R',\omega)\, A(s-u,R',\omega)\, du\, dv}{1 - \int_V \int_{-\infty}^{\infty} \frac{s^2 \exp\left[\frac{iR'}{c} (s \cos \psi' - u)\right] A(s-u,R',\omega)\, du\, dv}{4\pi R'}} \,. \tag{4.20}$$

Substituting (4.20) into (4.19), the solution of the degenerate integral equation (4.18) becomes:

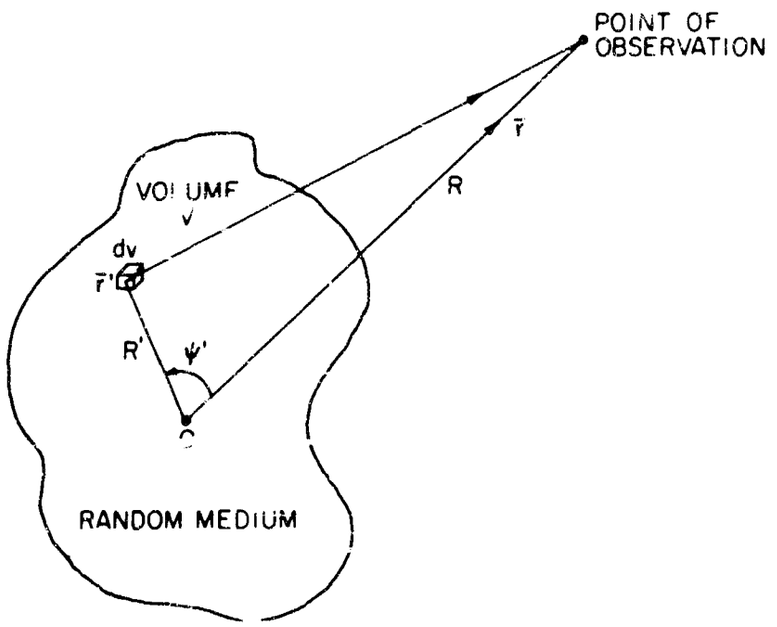


Figure 4.1 - Geometry for the Degenerate Kernel Approximation

$$\Gamma(s,\psi,\omega) = F(s,\psi,\omega)$$

$$+ \frac{s^2}{4\pi R} \frac{\int_V \int_{-\infty}^{\infty} \exp[-i \frac{s}{c} (R - R' \cos \psi')] F(u,R',\omega) A(s-u,R',\omega) \, du \, dv}{1 - \int_V \int_{-\infty}^{\infty} \frac{u^2}{4\pi R'} \exp[\frac{iR'}{c}(s \cos \psi' - u)] A(s-u,R',\omega) \, du \, dv} \, . \tag{4.21}$$

As in section 3.3, the resolvent kernel of the degenerate integral equation (4.18) can be defined from (4.21):

$$\Gamma(s,u,R,R',\omega) =$$

$$\frac{s^2}{4\pi R} \frac{\exp[-i \frac{s}{c} (R - R' \cos \psi')] A(s-u,R',\omega) \, du \, dv}{1 - \int_V \int_{-\infty}^{\infty} \frac{u^2}{4\pi R'} \exp[\frac{iR'}{c}(s \cos \psi' - u)] A(s-u,R',\omega) \, du \, dv} \, . \tag{4.22}$$

Inspection of (4.21) and (4.22) reveals the same difficulties that were discussed in section 3.3. Namely, the denominator of the resolvent kernel may vanish because $A(s-u,R',\omega)$ is a stochastic process. If the denominator vanishes, the resolvent kernel does not exist. However, it can be shown that this is an event of zero probability if the probability density function of γ,

$$\gamma = \int_V \int_{-\infty}^{\infty} \frac{u^2}{4\pi R'} \exp[\frac{iR'}{c}(s \cos \psi' - u)] A(s-u, R',\omega) \, du \, dv \tag{4.23}$$

exists and is continuous in some neighborhood of $\gamma = +1$. To compute the statistical measures of $Y(s,R,\omega)$, the denominator of the resolvent kernel must be expanded in a power series of γ. The power series expansion is valid with probability one if $|\gamma| < 1$ with probability one. Note that (4.23) does not have a singularity at $R' = 0$, because the volume element dv contains a R' squared term.

The power series expansion for (4.21) is:

$$Y(s,R,\omega) = F(s,R,\omega)$$

$$+ \frac{s^2}{4\pi R} \int_V \int_{-\infty}^{\infty} \exp\left[-i \frac{s}{c} (R - R' \cos \psi')\right] F(u,R',\omega)\, du\, dv$$

$$(1 + \gamma + \gamma^2 \cdots \gamma^n \cdots) \quad , \tag{4.24}$$

where γ is given by (4.23). It can be seen from (4.24) that the ensemble average of $Y(s,R,\omega)$ is zero if the ensemble average of $F(s,R,\omega)$ is zero. This follows from the assumption that $X(t,\bar{r},\omega)$ and $\alpha(t,\bar{r},\omega)$ are statistically independent, zero mean stochastic processes. The statistical measures that are of greater interest are the power spectral density and the coherence functions. These quantities are Fourier transform pairs of one another. The power spectrum $\Phi_{xx}(s)$ is related to the spectral representation of the random variable by*

$$< d\,X\,(s_1)\, d\,X^*\,(s_2) > = \delta(s_1 - s_2)\, \Phi_{xx}(s_1)\, d\,s_1\, d\,s_2 \quad . \tag{4.25}$$

---

*Discussed in greater detail in Chapter II, sections 2.6 and 2.7.

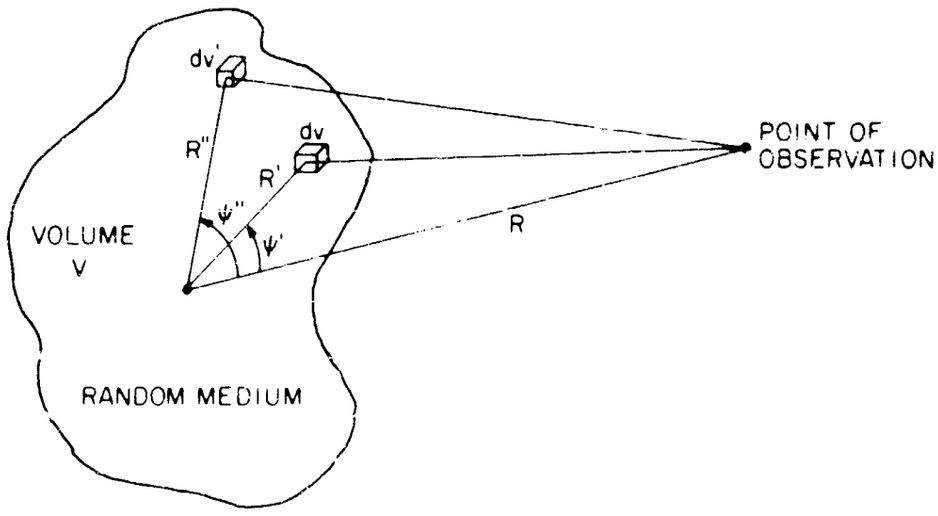


Figure 4.2 - Geometry for Double Integration

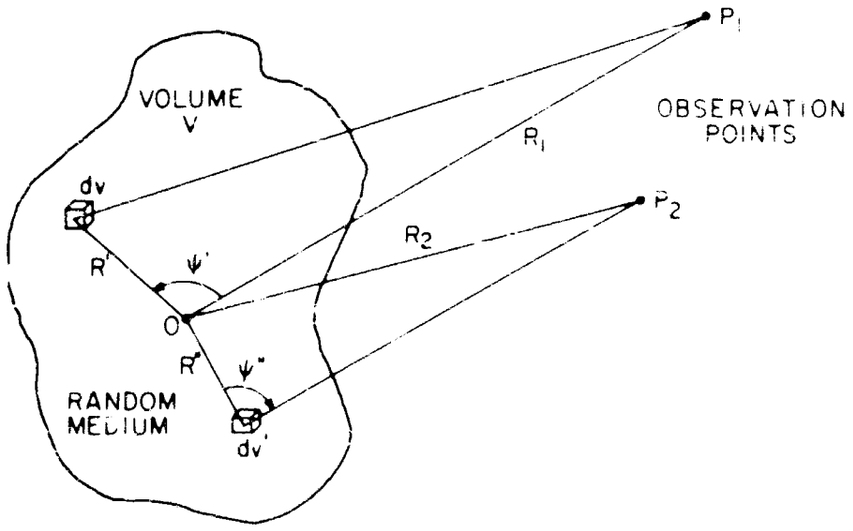


Figure 4.3 - Geometry for Computing Mutual Coherence Function

Because we are eventually interested in computing the coherence between two spatially separated points (Figure 4.3), we calculate $\langle Y(s_1, R_1, \omega)\, Y^*(s_2, R_2, \omega)\rangle$. Using just the first term in the power series expansion (4.24), we have:

$$\langle Y(s_1, R_1, \omega)\, Y^*(s_2, R_2, \omega)\rangle \approx \langle F(s_1, R_1, \omega)\, F^*(s_2, R_2, \omega)\rangle +$$

$$\frac{s_1^2\, s_2^2}{(4\pi)^2\, R_1 R_2} \int_V \int_V \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\left[-i\frac{s_1}{c}(R_1 - R'\cos\psi') + i\frac{s_2}{c}(R_2 - R''\cos\psi'')\right]$$

$$\langle F(u, R', \omega)\, F^*(u', R'', \omega)\rangle \langle A(s_1 - u, R', \omega)\, A^*(s_2 - u', R'', \omega)\rangle\, du\, du'\, dv\, dv'$$

$$= \Phi_{FF}(s_1, R_1, R_2)\, \delta(s_1 - s_2) +$$

$$s_1^2\, s_2^2 \int_V \int_V \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H(s_1, s_2, R_1, R_2, R', R'')\, \Phi_{FF}(u, R', R'')\, \delta(u - u')$$

$$\langle A(s_1 - u, R', \omega)\, A^*(s_2 - u', R'', \omega)\rangle\, du\, du'\, dv\, dv'$$

$$= \Phi_{FF}(s_1, R_1, R_2)\, \delta(s_1 - s_2) +$$

$$\left[s_1^2\, s_2^2 \int_V \int_V \int_{-\infty}^{\infty} H(s_1, s_2, R_1, R_2, R', R'')\, \Phi_{FF}(u, R', R'')\right.$$

$$\left.\Phi_{\alpha\alpha}(s_1 - u, R', R'')\, du\, dv\, dv'\right]\, \delta(s_1 - s_2)\ , \qquad (4.26a)$$

where

$$H(s_1,s_2,R_1,R_2,R',R'') = \frac{1}{(4\pi)^2 R_1 R_2} \exp[-i\frac{s_1}{c}(R_1-R'\cos\psi') + i\frac{s_2}{c}(R_2-R''\cos\psi'')] \quad . \tag{4.26b}$$

The cross terms of the form

$$< F(s_1,R_1,\omega)\frac{s_2^2}{4\pi R_2} \int_V \int_{-\infty}^{\infty} \exp[+i\frac{s}{c}(R - R'\cos\psi')]\, F^*(u,R',\omega)\, A^*(s_2-u,R',\omega) > du\, dv \tag{4.27}$$

vanish because the average $< A^*(s_2-u,R',\omega)>$ vanishes.

From (4.26), we note that using just the first term of the expansion, $< Y(s_1,R_1,\omega)\, Y^*(s_2,R_2,\omega) >$ is equal to a delta function times the sum of two terms. The coefficients of the delta function $\delta(s_1-s_2)$ are interpreted as the power spectral density of the solution of the scalar wave equation. The fact that we get a delta function $\delta(s_1-s_2)$ shows that, at least as far as the first approximation is concerned, the solution of the scalar wave equation is a wide-sense stationary process.* (For greater generality, the correlation between two spatially

*Discussed in section 2.

separated wave functions was computed. Of course, we have a special case when R_1 equals R_2.) Thus, the simplest approximation of the power spectral density of the wave function is, from (4.26),

$$\Phi_{yy}(s,R_1,R_2) \approx \Phi_{FF}(s,R_1,R_2) +$$

$$+ s^4 \int_V \int_V \int_{-\infty}^{\infty} H(s,s,R_1,R_2,R',R'') \, \Phi_{\alpha\alpha}(s-u,R',R'') \, \Phi_{FF}(u,R',R'') \, du \, dv \, dv'. \tag{4.28}$$

This crude approximation has already some interesting features. The first term is the power spectral density of the wave in the deterministic time-invariant medium. The second term demonstrates the spreading of the power spectral density of the wave function by the randomly time-varying medium. Even if $\Phi_{\alpha\alpha}(s,R',R'')$ is a lowpass function and its power spectrum does not overlap the power spectrum of $\Phi_{FF}(u,R',R'')$, the power spectrum of $\Phi_{yy}(s,R_1,R_2)$ is modified by the randomly time-varying medium. This is illustrated on Figure 4.4. Thus, the widely used quasimonochromaticity assumption is incorrect, or at least should be seriously questioned. Equation (4.28) is an approximate expression for the stochastic Green's function. It is an integral expression that relates the power spectral density of the solution to the spectral density of the source term and stochastic medium.

The next approximation can be computed by including the γ term in the power series expansion (4.24).

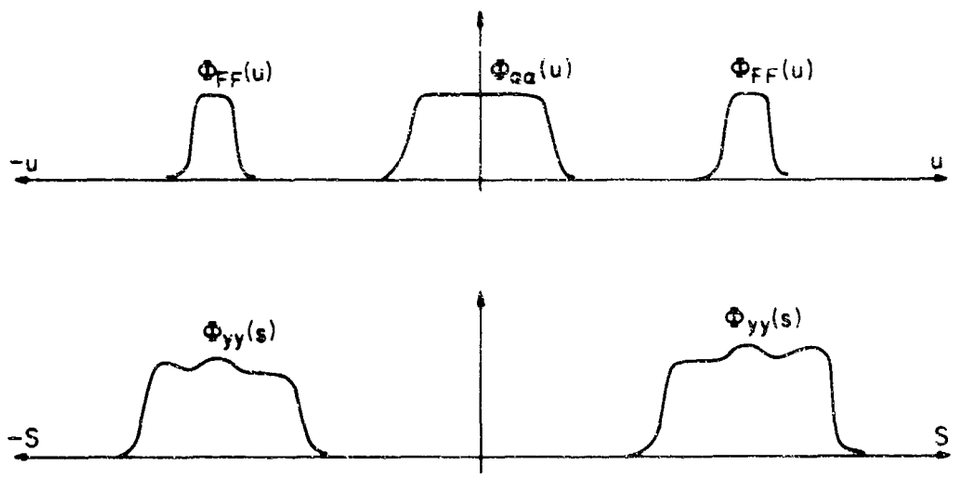


Figure 4.4 - Spreading of Power Spectrum by Randomly Time-Varying Medium

The cross terms give:

$$< F^{*}(s_2,R_2,\omega)\ \frac{s_1^2}{4\pi R_1}\ [\ \int_V \int_{-\infty}^{\infty} \exp\ [-i\ \frac{s_1}{c}\ (R_1-R'\cos\psi')]\ F(u,R',\omega)$$

$$A(s_1-u,R',\omega)\ du\ dv\ ]\ \gamma\ (s_1) > \ =$$

$$\frac{s_1^2}{4\pi R_1} \int_V \int_V \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{u'^2}{4\pi R''} \exp[-i\ \frac{s_1}{c}\ (R_1-R'\cos\psi')+ \frac{iR''}{c}(s_1\cos\psi''-u')]$$

$$< F(u,R',\omega)\ F^{*}(s_2,R_2,\omega)> < A(s_1-u,R',\omega)\ A(s_1-u',R'',\omega)>\ du\ du'\ dv\ dv'$$

$$= \frac{s_1^2}{4\pi R_1} \int_V \int_V \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{u'^2}{4\pi R''} \exp\ [-\frac{is_1}{c}\ (R_1-R'\cos\psi')+\frac{iR''}{c}\ (s_1\cos\psi''-u')]$$

$$\Phi_{FF}(u,R',R_2)\ \delta(u-s_2)\ \Phi_{\alpha\alpha}(s_1-u,R',R'')\ \delta(u+u')\ dv\ dv'\ du\ du'$$

$$= s_1^2\ s_2^2 \int_V \int_V \frac{1}{(4\pi)^2 R_1 R''} \exp[-\frac{is_1}{c}(R_1-R'\cos\psi')+ \frac{iR''}{c}(s_1\cos\psi'' + s_2)]$$

$$\Phi_{FF}(s_1,R',R_2)\ \Phi_{\alpha\alpha}(s_1+s_2,R',R'')\ dv\ dv'\ ,$$

(4.29)

and

$$< (F(s_1,R_1,\omega) \frac{s_2^2}{4\pi R_2} \int_V \int_{-\infty}^{\infty} \exp\left[i \frac{s_2}{c} (R_2 - R'\cos\psi')\right] F^*(u,R',\omega)$$

$$A^*(s_2-u,R',\omega)\, du\, dv)\, \gamma^*(s_2) > =$$

$$\frac{s_2^2}{4\pi R_2} \int_V \int_V \int_{-\infty}^{\infty} \frac{u'^2}{4\pi R''} \exp\left[i \frac{s_2}{c} (R_2 - R'\cos\psi') - \frac{iR''}{c} (s_2\cos\psi'' - u')\right]$$

$$< F(s_1,R_1,\omega)\, F^*(u,R',\omega) > < A^*(s_2-u,R',\omega)\, A^*(s_2-u',R'',\omega) > du\, du'\, dv\, dv'$$

$$= \frac{s_2^2}{4\pi R_2} \int_V \int_V \int_{-\infty}^{\infty} \frac{u'^2}{4\pi R''} \exp\left[i \frac{s_2}{c} (R_2 - R'\cos\psi') - \frac{iR''}{c} (s_2\cos\psi'' - u')\right]$$

$$\Phi_{FF}(u,R_1,R')\, \delta(u-s_1)\, \Phi_{\alpha\alpha}(s_2-u',R',R'')\, \delta(u+u')\, du\, du'\, dv\, dv'$$

$$= s_1^2 s_2^2 \int_V \int_V \frac{1}{(4\pi)^2 R_2 R''} \exp\left[\frac{is_2}{c} (R_2 - R'\cos\psi') - \frac{iR''}{c} (s_2\cos\psi'' + s_1)\right]$$

$$\Phi_{FF}(s_1,R_1,R')\, \Phi_{\alpha\alpha}(s_1+s_2,\, R',R'')\, dv\, dv' \quad .$$

(4.30)

These two terms exhibit the nonstationary behavior of the solution, because of the absence of delta functions $\delta(s_1-s_2)$. However, these terms vanish if the power spectra $\Phi_{\alpha\alpha}(s_1+s_2)$ and $\Phi_{FF}(s_1)$ do not overlap as happens in some cases of practical interest. If this is true for all s_1 and s_2, then (4.29) and (4.30) vanish and the wave function still remains wide sense stationary. Note that even in the case of the nonoverlapping power spectra, the time-varying random medium still spreads the power spectrum of the wave function as shown by (4.28).

The next term that contains a delta function $\delta(s_1-s_2)$ is:

$$s_1^2\, s_2^2 \int_V \int_V \int_{-\infty}^{\infty} \int H(s_1,s_2,R_1,R_2,R',R'') < F(u,R',\omega)\, F^*(u',R'',\omega) >$$

$$< A(s_1-u,R<,\omega)\, A^*(s_2-u',R'',\omega)\, \gamma(s_1)\, \gamma^*(s_2) >\, du\, du'\, dv\, dv' \quad . \tag{4.31}$$

To evaluate this term, we need to know the fourth moment of $A(u,R,\omega)$ or make an additional assumption that $\alpha(t,R,\omega)$ is Gaussian. If $\alpha(t,R,\omega)$ is Gaussian, then $A(u, R,\omega)$ and γ are also Gaussian processes.* This, however, violates the condition for validity of the power series expansion of the denominator of (4.21). For the power series (4.24) to converge almost everywhere in probability, $|\gamma| < 1$ a.e.p. This condition is not satisfied when γ is a Gaussian process, and the power series does not converge. This also follows from the actual

*Discussed in section 2.8 and also in Moyal (1949).

computation of the higher order terms in the power series expansion. For example, (4.31) produces a sum of three averages or, in general computation of the average of 2n A(u, ∧,ω)'s, gives a sum of $\frac{(2n)!}{2^n\, n!}$ terms. Because of this large number of terms except for the trivial case of zero variance, the power series expansion diverges when γ is a Gaussian random process. In such a case, the power series expansion is not useful and other methods for computing the statistical measures of (4.21) must be found.

Since (4.22) is a nonlinear function of $A(s-u,R',\omega)$, the various techniques for calculation of statistical measures at the output of nonlinear devices can be used to calculate the statistical measures of (4.22) (Middleton, 1960; and Deutsch, 1962). The nonlinear transformation of random processes is a specialized topic in itself, and it is not discussed any further here.

Simplifying (4.31) and retaining only the coefficient of $\delta(s_1-s_2)$, the next approximation to the stationary part of the power spectral density $\Phi_{yy}(s,R_1,R_2)$ becomes, for the Gaussian coefficient*

$$\Phi_{yy}(s,R_1,R_2) \approx \Phi_{FF}(s,R_1,R_2) +$$

$$+ \varepsilon^4 \int_V \int_V \int_{-\infty}^{\infty} H(s,s,R_1,R_2,R',R'')\, \Phi_{\alpha\alpha}(s-u,R',R'')\, \Phi_{FF}(u,R',R'')\, du\, dv\, dv'$$

*As stated before, the power series expansion (4.24) does not converge for the Gaussian coefficient. This step is for illustration purposes only.

$$(1 + \int_V \int_V \int_{-\infty}^{\infty} H_1(s,u,R',R'')\, \Phi_{\alpha\alpha}(s-u,R',R'')\, du\, dv\, dv' \ldots),$$

(4.32a)

where $H(s,s,R_1,R_2,R',R'')$ is defined by (4.28b) and

$$H_1(s,u,R',R'') =$$

$$\frac{u^4}{(4\pi)^2 R'R''} \exp\left[\frac{is}{c}(R'\cos\psi' - R''\cos\psi'') - \frac{iu}{c}(R' - R'')\right] .$$

(4.32b)

Thus, beside the stationary term, two nonstationary terms are obtained, even when the random coefficient $\alpha(t,R,\omega)$ is a Gaussian process. If $\alpha(t,R,\omega)$ is not a Gaussian process, as mentioned before, knowledge of its fourth moment is required for the detailed evaluation of (4.31). If more terms are included in the power series expansion, knowledge of still higher order statistics of $\alpha(t,R,\omega)$ is required. Even if the power series (4.24) converges almost everywhere, this does not imply the existence of $< Y(s_1,R_1,\omega)\, Y^*(s_2,R_2,\omega) >$ because $Y(s_1,R_1,\omega)$ and $Y^*(s_2,R_2,\omega)$ may not be second-order random processes; that is, $< |Y(s_1,R_1,\omega)|^2 >$ and $< |Y(s_2,R_2,\omega)|^2 >$ may not be finite. (It is perfectly reasonable for the power spectral density to have singularities provided they are integrable.) In such case, one may calculate the integrated power spectrum, denoted by $S'(s',s'',R_1,R_2)$

by integrating $\Phi_{yy}(s,R_1,R_2)$ over the frequency band from s' to s":

$$S(s',s'',R_1,R_2) = \int_{s'}^{s''} \Phi_{yy}(s,R_1,R_2)\, ds \quad . \tag{4.33}$$

The above integral is only applicable for wide-sense stationary stochastic processes. If one takes into account the nonstationary part of the solution, the power spectral density is a function of two frequencies as discussed in section 2.7. Hence, the integrated power spectrum must be written as a double integral:

$$S'(s',s'',s''',s'''',\ R_1,R_2) = \int_{s'}^{s''}\int_{s'''}^{s''''} < Y(s_1,R_1,\omega)\ Y^{*}(s_2,R_2,\omega) > ds_1\ ds_2 \ . \tag{4.34}$$

The integrated power spectrum is actually a better representation of the physically measured energy, because all instruments have nonzero bandwidth.

Other statistical measures of interest are the mutual coherence function and normalized mutual coherence function (Born and Wolf, 1964; Beran and Parrent, 1964; O'Neill, 1963; and Mandel and Wolf, 1965). We denote these by $C_{yy}(R_1,R_2,\tau)$ and by $c_{yy}(R_1,R_2,\tau)$ for the wide-sense stationary processes, and by $C_{yy}(R_1,R_2,t_1,t_2)$ and $c_{yy}(R_1,R_2,t_1,t_2)$ for the nonstationary processes where, in the first case, $\tau = t_2 - t_1$. The symbols Γ and γ are used in the literature for the coherence function and the normalized coherence function, but Γ has already been used to designate the resolvent kernel of the integral equation. γ has

also been frequently used. For the wide-sense stationary process, the mutual coherence function $C_{yy}(R_1,R_2,\tau)$ is simply the Fourier transform of $\Phi_{yy}(s,R_1,R_2)$:

$$C_{yy}(R_1,R_2,\tau) = \int_{-\infty}^{\infty} \Phi_{yy}(s,R_1,R_2)\, e^{is\tau}\, ds \quad , \tag{4.35}$$

and the normalized mutual coherence function $c_{yy}(R_1,R_2,\tau)$ is

$$c_{yy}(R_1,R_2,\tau) = \frac{C_{yy}(R_1,R_2,\tau)}{[C_{yy}(R_1,R_1,0)\; C_{yy}(R_2,R_2,0)]^{\frac{1}{2}}} =$$

$$= \frac{\int_{-\infty}^{\infty} \Phi_{yy}(s,R_1,R_2)\; e^{is\tau}\; ds}{\left| \int_{-\infty}^{\infty} \Phi_{yy}(s,R_1,R_1)\, ds \int_{-\infty}^{\infty} \Phi_{yy}(s,R_2,R_2)\, ds \right|^{\frac{1}{2}}} \quad . \tag{4.36}$$

c_{yy} is also known as the complex degree of coherence. It is a useful concept in interpreting interference patterns of partially coherent light (Beran and Parrent, 1964). The concept is also useful in optical and radio astronomy. Mandel and Wolf (1965) show that, on one hand $c_{yy}(R_1,R_2,\tau)$ is a measure of the correlation of the complex field at two points R_1 and R_2, and, on the other hand, it is a measure of the sharpness and location of the fringe maxima obtained

by superposing the beams propagated from these points. $C_{yy}(R_1,R_2,\tau)$ and $c_{yy}(R_1,R_2,\tau)$ here are ensemble coherence functions because they are obtained by ensemble averaging, rather than time averaging. Both ensemble and time averaged coherence functions are used (Mandel and Wolf, 1965). From Schwarz's inequality, we have:

$$0 \leq |c_{yy}(R_1,R_2,\tau)| \leq \frac{|C_{yy}(R_1,R_2,\tau)|}{[C_{yy}(R_1,R_1,0)]^{\frac{1}{2}}\,[C_{yy}(R_2,R_2,0)]^{\frac{1}{2}}} \leq 1 \quad . \tag{4.37}$$

These extremes characterize complete incoherence and coherence, respectively. The quantities $C_{yy}(R_1,R_1,\tau)$ and $C_{yy}(R_2,R_2,\tau)$ are self coherence functions at R_1 and R_2. $C_{yy}(R_1,R_1,0)$ and $C_{yy}(R_2,R_2,0)$ are the ordinary light intensity (in optics) at R_1 and R_2.

When the wave function is not wide-sense stationary, the coherence functions must be defined in terms of the bifrequency Fourier transforms;* that is,

$$C_{yy}(R_1,R_2,t_1,t_2) =$$

$$= \int_{-\infty}^{\infty}\int \langle Y(s_1,R_1,\omega)\, Y^*(s_2,R_2,\omega)\rangle\, e^{-is_1t_1 + is_2t_2}\, ds_1\, ds_2 \quad , \tag{4.38a}$$

*Discussed in section 2.7.

$$C_{yy}(R_1,R_1,t_1,t_1) =$$

$$= \int_{-\infty}^{\infty}\int < Y(s_1,R_1,\omega)\ Y^*(s_2,R_1,\omega) > e^{-is_1t_1 + is_2t_2}\ ds_1\ ds_2 \ ,$$

(4.38b)

and

$$C_{yy}(R_2,R_2,t_2,t_2) =$$

$$= \int_{-\infty}^{\infty}\int < Y(s_1,R_2\ \omega)\ Y^*(s_2,R_2,\omega) > e^{-is_1t_2 + is_2t_2}\ ds_1\ ds_2 \ .$$

(4.38b)

and

$$C_{yy}(R_2,R_2,t_2,t_2) =$$

$$= \int_{-\infty}^{\infty}\int < Y(s_1,R_2,\omega)\ Y^*(s_2,R_2,\omega) > e^{-is_1t_2 + is_2t_2}\ ds_1\ ds_2 \ .$$

(4.38c)

The normalized mutual coherence function $c_{yy}(R_1,R_2,t_1,t_2)$ can be defined for a wave function that is not wide-sense stationary by:

$$c_{yy}(R_1,R_2,t_1,t_2) = \frac{C_{yy}(R_1,R_2,t_1,t_2)}{\left[\,|C_{yy}(R_1,R_1,t_1,t_1)|\;|C_{yy}(R_2,R_2,t_2,t_2)|\,\right]^{\frac{1}{2}}} \tag{4.39}$$

The normalized mutual coherence function as defined by (4.39) is a generalization of the wide-sense stationary concepts to the case when the wave function is not wide-sense stationary. The Fourier transforms (4.35) and (4.38) express the mutual coherence functions for the wide-sense stationary and for the nonstationary processes. If satisfactory approximations for the $< Y(s_1,R_1,\omega)\;\; Y^*(s_2,R_2,\omega) >$ have been obtained, then, at least in principle, it is a simple matter to compute the mutual coherence functions from (4.35) or (4.38). To carry out the detailed computation of these Fourier transforms, knowledge of the statistics of the source and stochastic medium is needed. In particular, for the source, we need the power spectral density $\Phi_{FF}(u)$, and for the stochastic coefficient, we need its power spectral density and higher order spectral moments. We have indicated a method by which the mutual coherence functions may be calculated for a scalar wave propagating in a randomly time- and space-varying medium. According to Beran and Parrent (1964), this has been an unsolved problem.

In the next section, the Neumann series expansion of the integral equation (4.16) is investigated.

4.4 Neumann Series Expansion

To obtain a Neumann series solution for the integral equation (4.16), a solution of the following form is assumed:

$$Y(s,\bar{r},\omega) = \lambda^0 Y_0(s,\bar{r},\omega) + \lambda^1 Y_1(s,\bar{r},\omega) + \ldots \lambda^n Y_n(s,\bar{r},\omega) \ldots \quad (4.40)$$

Substituting the assumed solution (4.40) into the integral equation

$$Y(s,\bar{r},\omega) = F(s,\bar{r},\omega) + \lambda L^{-1} s^2 \int_{-\infty}^{\infty} Y(u,\bar{r}',\omega) A(s-u,\bar{r}',\omega)\, du \quad , \quad (4.41)$$

and equating the coefficients of the same powers of λ, the following iterative solution is obtained:

$$Y_0(s,\bar{r},\omega) = F(s,\bar{r},\omega) \quad , \quad (4.42a)$$

$$Y_1(s,\bar{r},\omega) = L^{-1} s^2 \int_{-\infty}^{\infty} Y_0(u,\bar{r}',\omega) A(s-u,\bar{r}',\omega)\, du$$

$$= L^{-1} s^2 \int_{-\infty}^{\infty} F(u,\bar{r}',\omega) A(s-u,\bar{r}',\omega)\, du \quad , \quad (4.42b)$$

and

$$Y_2(s,\bar{r},\omega) = L^{-1}\, s^2 \int_{-\infty}^{\infty} Y_1(u,\bar{r}',\omega)\, A(s-u,\bar{r}',\omega)\, du$$

$$= L^{-1}\, s^2 \int_{-\infty}^{\infty} L^{-1}\, u^2\, du \int_{-\infty}^{\infty} F(u',\bar{r}'',\omega)\, A(u-u',\bar{r}',\omega)\, du'. \tag{4.42c}$$

The general term is

$$Y_n(s,\bar{r},\omega) = L^{-1}\, s^2 \int_{-\infty}^{\infty} Y_{n-1}(u,\bar{r}',\omega)\, A(s-u,\bar{r}',\omega)\, du \quad , \tag{4.42d}$$

where L^{-1} is the inverse operator for the Helmholtz equation and its boundary conditions. In the spherical coordinate system, the inverse operator for an outgoing wave is:

$$-\int_V dv\, \frac{\exp\left[-i\frac{s}{c}\,|\,\bar{r}-\bar{r}'\,|\right]}{4\pi\,|\,\bar{r}-\bar{r}'\,|}\,(\cdot) \quad , \tag{4.43}$$

where dv is the volume element in the spherical coordinate system, and $\int_V$ indicates volume integration over the space containing the randomly space- and time-varying medium. For simplicity, denote

$$G(s,\bar{r},\bar{r}') = -\frac{\exp\left[-i\frac{s}{c}\,|\,\bar{r}-\bar{r}'\,|\right]}{4\pi\,|\,\bar{r}-\bar{r}'\,|} \quad . \tag{4.44}$$

Recalling from (4.16) that $\lambda = -1$, the Neumann series iteration becomes:

$$\begin{aligned} Y(s,\bar{r},\omega) = {} & F(s,\bar{r},\omega) + \\ & + \lambda s^2 \int_V \int_{-\infty}^{\infty} G(s,\bar{r},\bar{r}')\, A(s-u,\bar{r}',\omega)\, F(u,\bar{r},\omega)\, dv\, du \\ & + \lambda^2 s^2 \int_V \int_V \int_{-\infty}^{\infty} \int (u^2 G(s,\bar{r},\bar{r}') G(u,\bar{r}',\bar{r}'') \\ & A(s-u,\bar{r}')\, A(u-\bar{u}'',\bar{r}'')\, F(u',\bar{r}'',\omega))\, dv\, dv'\, du\, du' \\ & + \quad . \\ & \quad . \\ & \quad . \end{aligned} \tag{4.45}$$

This series has some interesting physical interpretations. The first term after $F(s,\bar{r},\omega)$ can be interpreted as the first-order scattering of the wave by the random medium. It is the widely used Born approximation. A number of papers and books consider only the Born approximation.

The Born approximation gives a satisfactory answer if only single-scattering is important. The higher order terms are the multiple-scattering corrections and must be taken into account in long distance wave propagation through the random medium.

The single-scattering approximation is investigated first. After that, the multiple scattering and the convergence of the Neumann series expansion is investigated. The statistical measures of interest are the power spectral density and the mutual coherence functions. The single scattering approximation gives:

$$< Y(s_1,\bar{r}_1,\omega)\ Y^*(s_2,\bar{r}_2,\omega)> \approx < F(s_1,\bar{r}_1,\omega)\ F^*(s_2,\bar{r}_2,\omega)>$$

$$+ \lambda^2 s_1^2 s_2^2 \int_V \int_V \int_{-\infty}^{\infty} \int G(s_1,\bar{r}_1,\bar{r}') G^*(s_2,\bar{r}_2,\bar{r}'') < F(u,\bar{r}',\omega)\ F(u',\bar{r}'',\omega) >$$

$$< A(s_1-u,\bar{r}',\omega)\ A^*(s_2-u',\bar{r}'',\omega) >\ du\ du'\ dv\ dv'$$

$$= \Phi_{FF}(s_1,\bar{r}_1,\bar{r}_2)\ \delta(s_1-s_2)\ +$$

$$\lambda^2 s_1^2 s_2^2 \int_V \int_V \int_{-\infty}^{\infty} \int G(s_1,\bar{r}_1,\bar{r}')\ G^*(s_2,\bar{r}_2,\bar{r}'')\ \Phi_{FF}(u,\bar{r}',\bar{r}'')\ \delta(u-u')$$

$$< A(s_1-u,\bar{r}',\omega)\ A^*(s_2-u',\bar{r}'',\omega)>\ du\ du'\ dv\ dv'$$

$$= \Phi_{FF}(s_1,\bar{r}_1,\bar{r}_2)\,\delta(s_1-s_2)$$

$$+ \left[\lambda^2 s_1^2 s_2^2 \int_V \int_V G(s_1,\bar{r}_1,\bar{r}')\, G^*(s_2,\bar{r}_2,\bar{r}'') \right.$$

$$\left. \Phi_{\alpha\alpha}(s_1-u,\bar{r}',\bar{r}'')\, \Phi_{FF}(u,\bar{r}',\bar{r}'')\,du\,dv\,dv' \right] \delta(s_1-s_2) \quad . \tag{4.46}$$

The cross terms have vanished because $A(s-u,\bar{r},\omega)$ is a process with zero mean. The approximate expression for the power spectral density is the sum of the coefficients of the delta function $\delta(s_1-s_2)$; i.e.,

$$\Phi_{yy}(s,\bar{r}_1,\bar{r}_2) \approx \Phi_{FF}(s,\bar{r}_1,\bar{r}_2) +$$

$$s^4 \int_V \int_V \int_{-\infty}^{\infty} G(s,\bar{r}_1,\bar{r}')G^*(s,\bar{r}_2,\bar{r}'')\, \Phi_{\alpha\alpha}(s-u,\bar{r}',\bar{r}'')\, \Phi_{FF}(u,\bar{r}',\bar{r}'')\, du\,dv\,dv' \quad . \tag{4.47}$$

This expression reduces to (4.28) if the approximations (4.17) are used to simplify the factor $G(s,\bar{r}_1,\bar{r}')\, G^*(s_1,\bar{r}_2,\bar{r}'')$; that is, this factor becomes then $H(s,s,R_1,R_2,R',R')$ as defined by (4.26b). This expression again demonstrates the spreading of the power spectral density of the wave function by the randomly time-varying medium. Equation (4.47) again shows that, as far as the first approximation

is concerned, the wave function is a wide-sense stationary process. An approximation for the normalized mutual coherence function can be obtained by substituting (4.47) into

$$c_{yy}(\bar{r}_1,\bar{r}_2,\tau) = \frac{\int_{-\infty}^{\infty} \Phi_{yy}(s,\bar{r}_1,\bar{r}_2)\, e^{-is\tau}\, ds}{\left[\int_{-\infty}^{\infty} \Phi_{yy}(s,\bar{r}_1,\bar{r}_2)\, ds \int_{-\infty}^{\infty} \Phi_{yy}(s,\bar{r}_2,\bar{r}_2)\, ds\right]^{\frac{1}{2}}} \tag{4.48}$$

and carrying out the indicated integration.

If multiple scattering is important, more terms must be included in the Neumann series expansion. In such instances, we have for the n^{th} approximation

$$\begin{aligned} &\langle Y(s_1,\bar{r}_1,\omega)\, Y^*(s_2,\bar{r}_2,\omega)\rangle_n = \\ &= \langle (\lambda^0 Y_0(s_1,\bar{r}_1,\omega) + \lambda Y_1(s_1,\bar{r}_1,\omega) + \lambda^2 Y_2(s_1,\bar{r}_1,\omega) + \cdots \\ &\quad + \lambda^n Y_n(s_1,\bar{r}_1,\omega))\,(\lambda^0 Y_0^*(s_2,\bar{r}_2,\omega) + \lambda^1 Y_1^*(s_2,\bar{r}_2,\omega) + \\ &\quad \lambda^2 Y_2^*(s_2,\bar{r}_2,\omega) \cdots \lambda^n Y_n^*(s_2,\bar{r}_2,\omega)) \rangle \quad . \end{aligned} \tag{4.49}$$

Carrying out the indicated multiplication and averaging we get:

$$\langle Y(s_1,\bar{r}_1,\omega)\, Y^*(s_2,\bar{r}_2,\omega)\rangle_n =$$

$$= \lambda^0 \langle Y_0(s_1,\bar{r}_1,\omega)\, Y_0^*(s_2,\bar{r}_2,\omega)\rangle$$

$$+ \lambda^2 \{ \langle Y_0(s_1,\bar{r}_1,\omega)\, Y_2^*(s_2,\bar{r}_2,\omega)\rangle + \langle Y_2(s_1,\bar{r}_1,\omega)\, Y_0^*(s_2,\bar{r}_2,\omega)\rangle$$

$$+ \langle Y_1(s_1,\bar{r}_1,\omega)\, Y_1^*(s_2,\bar{r}_2,\omega)\rangle \} + \lambda^4 \{ \langle Y_2(s_1,\bar{r}_1,\omega)\, Y_2^*(s_2,\bar{r}_2,\omega)\rangle$$

$$+ \langle Y_3(s_1,\bar{r}_1,\omega)\, Y_1^*(s_2,\bar{r}_2,\omega)\rangle + \langle Y_1(s_1,\bar{r}_1,\omega)\, Y_3^*(s_2,\bar{r}_2,\omega)\rangle$$

$$+ \langle Y_0(s_1,\bar{r}_1,\omega)\, Y_4^*(s_2,\bar{r}_2,\omega)\rangle + \langle Y_4(s_1,\bar{r}_1,\omega)\, Y_0(s_2,\bar{r}_2,\omega)\rangle \}$$

$$+ \ldots \lambda^{2n} \text{ (2n + 1 terms)} \quad . \tag{4.50}$$

The coefficients of odd powers of λ are zero because it is assumed that the odd moments of $A(u,\bar{r},\omega)$. If this assumption is not made, the odd terms must be included. The terms $\langle Y_0(s_1,\bar{r}_1,\omega)\, Y_0^*(s_2,\bar{r}_2,\omega)\rangle$ and $\langle Y_1(s_1,\bar{r}_1,\omega)\, Y_1^*(s_2,\bar{r}_2,\omega)\rangle$ and $\langle Y_1(s_1,\bar{r}_1,\omega)\, Y_1^*(s_2,\bar{r}_2,\omega)\rangle$ were already computed for the first-order approximation (3.47). The λ^2 terms are:

$$< Y_o(s_1,\bar{r}_1,\omega) \quad Y_2^*(s_2,\bar{r}_2,\omega) > =$$

$$s_2^2 \int_V \int_V \int_{-\infty}^{\infty} \int u^2 G^*(s_2,\bar{r}_2,\bar{r}') G^*(u,\bar{r}',\bar{r}'') < A^*(s_2-u,\bar{r}') \; A^*(u-u',\bar{r}'') >$$

$$< F^*(u',\bar{r}'',\omega) \; F(s_1,\bar{r}_1,\omega) > \; du \; du' \; dv \; dv' =$$

$$s_2^2 \int_V \int_V \int_{-\infty}^{\infty} \int u^2 G^*(s_2,\bar{r}_2,\bar{r}') G^*(u,\bar{r}',\bar{r}'') < A^*(s_2-u,\bar{r}') \; A^*(u-u',\bar{r}'') >$$

$$\Phi_{FF}(u',\bar{r}_1,\bar{r}'') \; \delta(u'-s_1) \; du \; du' \; dv \; dv' =$$

$$s_2^2 \int_V \int_V \int_{-\infty}^{\infty} u^2 G^*(s_2,\bar{r}_2,\bar{r}') G^*(u,\bar{r}',\bar{r}'') < A^*(s_2-u,\bar{r}') A(s_1-u,\bar{r}'') >$$

$$\Phi_{FF}(s_1,\bar{r}',\bar{r}'') \; du \; dv \; dv' =$$

$$\delta(s_2-s_1) \; s_2^2 \int_V \int_V \int_{-\infty}^{\infty} u^2 G^*(s_2,\bar{r}_2,\bar{r}_1) \; G^*(u,\bar{r}',\bar{r}'')$$

$$\Phi_{\alpha\alpha}(s_2-u,\bar{r}'',\bar{r}') \; \Phi_{FF}(s_1,\bar{r}',\bar{r}'') \; du \; dv \; dv' \; . \qquad (4.51)$$

The contribution of this term to the power spectral density is

$$s^2 \int_V \int_V \int_{-\infty}^{\infty} u^2 G^*(s,\bar{r}_2,\bar{r}') G^*(u,\bar{r}',\bar{r}'')\, \Phi_{\alpha\alpha}(s-u,\bar{r}',\bar{r}'') \Phi_{FF}(s,\bar{r}'',\bar{r}_1)\, du\, dv\, dv' \quad . \tag{4.52}$$

Similarly, the contribution to the power spectral density from the term

$< Y_2(s_1,\bar{r},\omega)\, Y_o^*(s_2,r_2,\omega) >$ is:

$$s^2 \int_V \int_V \int_{-\infty}^{\infty} u^2 G(s,\bar{r}_1,\bar{r}') G(u,\bar{r}',\bar{r}'')\, \Phi_{\alpha\alpha}(s-u,\bar{r}',\bar{r}'')\, \Phi_{FF}(s,\bar{r}_2,\bar{r}'')\, du\, dv\, dv'. \tag{4.53}$$

These two terms, (4.52) and (4.53), do not contribute any energy to the frequencies which are outside the spectrum of $F(s,R,\omega)$ because, in both of these integrals, the argument of the power spectral density term ($\Phi_{FF}(s,\bar{r}'',\bar{r}_1)$ in (4.52) and $\Phi_{FF}(s,\bar{r}_2,\bar{r}'')$ in (4.53)) is s; i.e., same argument as the argument of the dependent variable. For both cases, the variable of integration is u. Therefore, at frequencies where the factors $\Phi_{FF}(s,\bar{r}'',\bar{r}_1)$ and $\Phi_{FF}(s,\bar{r}_2,\bar{r}'')$ in the integrand vanish, the respective integrals vanish and these terms do not contribute to the spreading of the power spectral density. This is not the case with the last term of (4.47) where the argument of

$\Phi_{FF}(u,\bar{r}',\bar{r}'')$ is same as the variable of integration (u). Hence, this term shows spectral spreading beyond the original band of frequencies. These two terms, (4.52) and (4.53), are also expressions for spectral density of wide-sense stationary processes because they, too, are coefficients of delta functions $\delta(s_1-s_2)$.

The computation of the higher order terms requires, in general, knowledge of higher order moments of the random coefficient $\alpha(t,R,\omega)$. Only if $\alpha(t,R,\omega)$ is a Gaussian process is knowledge of the first two moments sufficient to specify all higher moments. As shown in (4.50), λ^{2n} is multiplied by the sum of 2n + 1 averages. In the case of the Gaussian process, each average produces $\frac{(2n)!}{2^n\, n!}$ terms. For this reason, it can be expected that the Neumann series expansion diverges for the Gaussian processes. Detailed calculation shows that only one of the $\frac{(2n)!}{2^n\, n!}$ terms produces a delta function $\delta(s_1-s_2)$. Thus, the stationary part of the Neumann series expansion may converge (if additional conditions for convergence are satisfied) even if the nonstationary part of the series diverges.

The series expansion (4.50) converges if the magnitudes of the coefficients of λ^{2n} are smaller than $(2n + 1)\, M^{2n}$, where $M < 1$. If these conditions are true, then the series expansion (4.50) converges because it is dominated by

$$(1 - x)^{-2} = 1 + 2x + 3x^2 + \ldots (n + 1)\, x^n \ . \tag{4.54}$$

These are sufficient conditions for convergence; the series (4.50) may converge under less restrictive conditions. The u^2 factor in the

integrand of (4.52) does not cause any difficulty with the infinite integration because both $\Phi_{\gamma\gamma}(u)$ and $\Phi_{FF}(u)$ are either bandlimited or they behave in many cases approximately as $\frac{1}{u^4}$ or at least as $\frac{1}{u^2}$ for large u (Tatarski, 1961).

In the actual evaluation of the volume integrals, simplifying approximations can be made although we shall not do any such calculations. For example, in evaluation of the volume integral in (3.44), it can be observed that the main contribution to the integral comes from the volume where $\bar{r}' \approx \bar{r}''$; that is, from the volume where $|\bar{r}' - \bar{r}''|$ is of the same order as the correlation length of the spatial variation of the random medium. This is clearly so in the case of some of the widely used spatial correlation functions of the form:

$$b^2 \exp\left[-|\bar{r}' - \bar{r}''|/r_o\right] , \tag{4.55a}$$

$$b^2 \exp\left[-|\bar{r}' - \bar{r}''|^2/r_o^2\right] , \tag{4.55b}$$

and

$$\frac{b^2}{2^{\nu-1}\Gamma(\nu)} \left(\frac{|\bar{r}' - \bar{r}''|}{r_o}\right)^{\nu} K_{\nu}\left(\frac{|\bar{r}' - \bar{r}''|}{r_o}\right) , \tag{4.55c}$$

where r_o is the correlation length and $K_\nu(x)$ is the Bessel function of the second kind of imaginary argument (Tatarski, 1961). Booker and Gordon (1950) have used (4.55a) and, more recently, Krasilnikov*

*Krasilnikov's work is discussed by Tatarski (1961).

has used (4.55c). Tatarski (1961) discusses the theoretical justification for assuming these forms for the spatial correlation functions and compares the consequences of these assumptions with the experimental results. It has also been pointed out by Comstock (1963 and 1964) that the autocorrelation functions of the same form as (4.55c) are capable of satisfying the theoretical and experimental requirements for wave propagation problems in a random medium.

4.5 General Remarks and Conclusions

Both the Neumann series expansion and the degenerate kernel approximation demonstrate the spreading of the power spectrum of source by the time- and space-varying random medium. The quasimonochromatic assumption in the previous work has totally neglected this spreading of the power spectrum of the wave function. Even the first-order approximation provides a useful method for the computation of the approximate spreading of the power spectrum of the wave function. Higher order approximations show further spreading of the power spectrum. It thus appears that the successive convolutions of the power spectral density will spread the power spectral density of the wave function even further, make its spectrum wider and, therefore, its correlation time shorter. This agrees with the intuitive concept that the multiply scattered waves become decorrelated.

In both methods of solution, the first term in the solution (zeroth approximation) is the solution of the wave equation in the nonrandom medium. When the next approximation is used, the wave function is a wide-sense stationary stochastic process if the source and random medium are wide-sense stationary processes. If the higher

order approximations are used, terms appear which are no longer wide-sense stationary. For the first approximations, the mutual coherence function was a function of difference of the observation times; that is, $C(R_1,R_2,\tau)$, where $\tau = t_2-t_1$. But when higher order approximations are used, the mutual coherence function is a function of the actual observation times t_1 and t_2. The mutual coherence function must be written as $C(R_1,R_2,t_1,t_2)$ to express the dependence on t_1 and t_2. Because of the close relationship between the coherence functions and optical images, one would expect fluctuation of the light intensity when the light propagates through a randomly time-varying medium. Both "twinkling" and "quivering" are common in observation of stars by telescopes (Tatarski, 1961). Twinkling is the irregular fluctuation of the light intensity and quivering is the irregular fluctuation of the angle of arrival of the light. A large number of experimental papers have been devoted to the twinkling and quivering of stellar images. (Chapter 13 of Tatarski (1961) gives a number of references.) Under unfavorable observational conditions, instead of a luminous core and a series of concentric rings, one observes at the focal plane of a telescope an irregular patch of light which is dancing around and fluctuating in light intensity. A theoretical model which would explain this phenomenon could be based on the non-stationary mutual coherence functions.

Both the degenerate kernel approximation and the Neumann series expansion have the disadvantage that they are extremely laborious, but this is to be expected of a problem which is fundamentally very complex. In spite of this, even the first approximations provide useful results.

The Neumann series iterations represent multiple scattering of the incident wave. This interpretation appears to be useful even if the Neumann series expansion does not converge. What is needed is a method for interpreting the divergent part. W. P. Brown (1967) has used a selective summation method to interpret a divergent Neumann series expansion. The basic idea of the selective summation technique is to identify the terms in the Neumann series that causes the divergence, and to sum these terms to obtain a closed-form expression for the multiple scattering effects.

In the degenerate kernel method, we do have a closed form expression for the spectral representation of the wave function. However, when we attempt to compute the power spectral density and the coherence functions by expanding the denominator of the resolvent kernel in a power series, we may obtain a divergent power series. In this case, the power series expansion is not an essential part of the solution, it is just a computational method. If it does not work, other methods, such as the nonlinear random transformation methods (Middleton, 1960; and Deutsch, 1962) must be used. Application of the nonlinear random transformation techniques is complicated and it is a major topic in itself. Development of such computational methods would be a possible extension of this work. Other extensions of this work are discussed in the concluding chapter.

CHAPTER V

CONCLUSIONS, APPLICATIONS AND EXTENSIONS

In the diverse areas of electrical engineering, problems arise which should be properly described by linear differential equations with stochastic coefficients. In most cases, the randomness of the coefficients has been ignored because no widely applicable methods have been known for solving such problems. In this dissertation, an n^{th} order linear differential equation with such stochastic coefficients has been considered. It is assumed that the coefficients of the differential equation are separable into deterministic and stochastic parts. In the case of ordinary stochastic differential equations, the problem now becomes a problem of solving a Volterra integral equation with a stochastic kernel. Two methods of solution are considered: the Neumann series expansion and the degenerate kernel method. The Neumann series expansion is an expansion in terms of iterated integrals. A theorem which gives sufficient conditions for the uniform convergence of the Neumann series expansion is proved. The proof of this theorem and the actual Neumann series expansion is facilitated if the n^{th} order differential equation is expressed as n first-order differential equations. This formulation has the additional advantage of using the notation and terminology of the state-space formulation of modern control system theory. (This clarifies the connection between this work and the control system problems.)

The uniform convergence of the Neumann series expansion allows the solution of the stochastic differential equation to be expressed in terms of the resolvent kernel of the stochastic integral equation. The ensemble average and covariance function of the solution are expressed in terms of the corresponding statistical measures of the resolvent kernel and of the input process. The statistical measures of the resolvent kernel are functions of the Green's function of the deterministic operator and appropriate statistical measures of the stochastic coefficients. These results constitute a generalization of the corresponding expressions for linear time-varying systems to the linear randomly time-varying systems. The kernels of the integral expressions for the statistical measures of the solution can be interpreted as stochastic Green's functions.

The construction of the resolvent kernel by means of the Neumann series expansion is, however, extremely laborious. For this reason, other methods of solving the integral equation with the stochastic kernels were investigated. When the deterministic part of the stochastic differential operator is time-invariant, the separation of the stochastic and deterministic parts of the differential operator gives a Volterra integral equation with a degenerate kernel. In solving this Volterra integral equation, we still have to resort to Neumann series expansion. From the standpoint of computational difficulty and convergence of the solution, there is no essential difference between the straight Neumann series expansion and the degenerate kernel approach. Selection of the method depends on the physical problem and computational convenience. For example, in the case of a time-invariant deterministic part, the transition matrix may be

already known and, therefore, the degenerate kernel method may be convenient. In the case of the time-varying deterministic part, the straight Neumann series expansion is more convenient than the degenerate kernel method.

Thus, we have two complementary methods for solving stochastic Volterra integral equations. The degenerate kernel method is also important in that many arbitrary nondegenerate kernels can be approximated by degenerate kernels. Other approximate methods, such as the method of moments, are equivalent to the replacement of an arbitrary kernel with a degenerate kernel. Other methods, such as the Fredholm method and the Hilbert-Schmidt method do not offer any computational advantages for solving this problem. To use the Hilbert-Schmidt method with unsymmetrical kernels, one has to solve a pair of integral equations of the first kind. This problem is no less difficult than solving the integral equations of the second kind. The Fredholm theory has been very important in the development of the classical integral equation theory, but to use the Fredholm method for construction of the resolvent kernel is prohibitively difficult in practice.

Both the Neumann series expansion and the degenerate kernel method are very laborious. In practical computation, one must still resort to the truncation of the Neumann series expansions. The iterative procedure makes it possible to improve the approximations, because the truncation of the series is made almost as the last step in the solution. In other methods, such as in the hierarchy methods or in methods which use approximate differential equations, the approximations are made at the beginning of the problem. To improve these

approximations, one must essentially rework the whole problem. The method suggested here does not have this disadvantage.

It can be seen from the expressions for the covariance function of the solution that the knowledge of all the moments of the stochastic coefficients is required for the complete solution of the problem. If only the second-order statistics of the coefficients are known only an approximate solution of the problem is possible. This approximate solution can be obtained from the first-order approximation of the Neumann series expansion. An exceptional case is that in which the stochastic coefficients are Gaussian processes. Then, the knowledge of the second-order statistics is sufficient for the complete solution of the problem.

Both the Neumann series iteration and degenerate kernel approximation were applied to the investigation of the propagation of waves in a randomly space- and time-varying medium. Almost all the previous work has used the so-called quasimonochromatic assumption which essentially neglects the time variation of the medium. Such an assumption has been avoided in this dissertation. To solve the problem, all the stochastic quantities of the scalar wave equation are expressed by their spectral representation and the equation is solved for the spectral representation of the scalar wave function. From the spectral representation of the scalar wave function, its power spectral density and mutual coherence functions can be found. Both the Neumann series expansion and the degenerate kernel approximation demonstrate the spreading of the power spectrum of the source by the time-varying medium. In the Neumann series expansion, even the first-order approximation shows the spreading of the power spectrum. Higher order

approximations show further spreading of the power spectrum. The successive convolution of the power spectrum of the source term with the power spectrum of the random coefficient shows a further spread of the power spectral density of the wave function, making the correlation time of the wave function shorter. This agrees with the intuitive concept that multiply-scattered waves become decorrelated. The quasimonochromatic assumption does not show any spreading of the power spectrum.

Another interesting phenomenon is demonstrated by both methods of solution. The first term of the solution (zeroth approximation) is the solution of the wave equation in the non-random medium. When the next approximation is used, the wave function is a wide-sense stationary stochastic process (if the source and the stochastic medium are wide-sense stationary processes). If the higher order approximations are used, the solution contains, in addition to the wide-sense stationary terms, terms which are no longer wide-sense stationary. Thus the power spectral density must be expressed as a bifrequency Fourier transform of the mutual coherence function $C(R_1,R_2,t_1,t_2)$. The mutual coherence function is written in the above form to show that it is a function of the actual observation times, t_1 and t_2, not just the difference of the observation times. Because of the close relationship between the coherence function and optical images, one would expect fluctuation of the images which have been formed from light that is propagated through a randomly time-varying medium. Under unfavorable astronomical observation conditions, twinkling and quivering of the stellar images is indeed common. The existence of the nonstationary mutual coherence functions may be expected to be useful in

studying nonstationary optical images. Up to now, almost no work has been done in studying the properties of the mutual coherence functions in randomly time-varying media. The use of nonstationary coherence functions may be useful in the investigation of the factors which limit the performance of the optical and radio interferometers and high gain arrays.

An interesting future area of work is development of the statistical communication theory for multiplicative interference. Statistical communication theory has been based almost completely on the assumption that the interference has been added to the signal. This assumption is clearly not valid when one is working with rapidly fading signals or with signals that have been scattered by randomly time-varying media. Besides the usual additive noise, one has multiplicative noise. A useful approach to communication, radar or sonar system design would be to solve first the wave propagation problem through the stochastic medium. Then, the solution could be used to design the optimum signals and signal processor. The actual solution of the problem may be very complicated and, for tractability, one probably would have to base the analysis on the first approximation in the Neumann series expansion in the same manner as was done here.

There are a number of interesting applications of this work in control system theory. The obvious ones are the cases where the system parameters change randomly with time. For example, the center of gravity or the moment of inertia of a controlled vehicle changes due to motion of fuel in tanks or due to a change of operating conditions. The sensitivity analysis of control systems can also be based on stochastic differential equation theory. If the sensitivity of the

controlled variable to the change of a certain parameter is to be investigated, it may be assumed that this parameter is a random variable. Then, the stochastic differential equation is solved for a desired statistical measure of the controlled variable or error term. The solution is in terms of the Green's function (impulse response) of the system and the statistical measures of the random parameter. This functional relation constitutes a solution to the sensitivity problem. Incompletely identified control systems may also be treated as stochastic systems. In such cases, statistical measures of the system parameters are determined instead of a precise analytical expression for the coefficients.

In addition to the application of stochastic operator theory to control system and communication theory, it may be applied as a method for investigating the stochastic medium itself. For example, the sun has been used as a source of random signals to investigate the propagation of microwave signals through the atmosphere from the sun to a receiver on the earth. Measurements of this type are used to predict the propagation of signals between an orbiting communication satellite and a ground station. Mathematically, such a problem is completely analogous to the wave propagation problem considered here. Many other examples of this type can be cited. Among these would be investigation of plasmas by microwave and laser signals.

While a number of interesting results have been obtained in this dissertation, the work is far from being complete. Many useful and interesting extensions can be suggested. Specific problems should be worked out in great detail to check the practical utility of the computational methods. Problems should be realistically selected so

that the numerical results can be compared with the experimental results. An example of this is an electromagnetic wave propagation problem. The statistical measures of the medium's dielectric permittivity should be determined experimentally or derived theoretically from the theory of turbulence. Then, the wave equation should be solved for the statistical measure of interest, e.g., the mutual coherence function. Then, the mutual coherence function should be measured experimentally to check the ability of the theory to predict physical phenomenon. The various experimental methods for measuring the mutual coherence function are discussed in the IEEE special issue on partial coherence, (1967) and by Mandel and Wolf (1965) in their review paper.

Experimental comparisons are also possible in control system problems. A control systems problem could be solved by the methods of Chapter III, and the results could be compared with a hybrid computer simulation. The stochastic coefficients of the differential equations can be simulated by noise modulating the coefficient potentiometers of the operational amplifiers in the analogue portion of the computer. In such a setup, the analogue part of the computer solves the stochastic differential equation for a sample function (or realization) of the dependent variable. Computations are repeated for a large number of times and each time sample solutions are stored in the computer's memory. Then, the desired statistical measures, such as mean and correlation function, are computed by the digital portion of the computer from the stored sample solutions. Then the simulated results may be compared with the results obtained by either of the two methods in Chapter III. It may be expected that the mathematical methods developed in this dissertation would provide a systematic procedure for

understanding and interpretation of the results which are obtained in simulation of systems with stochastic parameters.

A number of generalizations and improvements to this work can be suggested. In the degenerate kernel method, it would be useful to find methods other than the power series expansion for computation of the statistical measures of the inverse matrix. Application of nonlinear transformation techniques may be useful in special cases. If tractable computational methods can be developed, the troublesome convergence problem of the power series expansion can be avoided. This would be especially valuable in the wave propagation problem. It is desirable to eliminate, in the wave propagation problem, the assumption that the dielectric permittivity is a wide-sense stationary stochastic process. A canonical integral expansion (Pugachev, 1965) of the stochastic processes should be used instead of the spectral representation, which is a special case of integral expansions for the wide-sense stationary stochastic processes. The solution with nonstationary coefficients can be expected to be more difficult than the wide-sense stationary case, since the familiar Fourier transform techniques are not directly applicable.

In Chapter IV, propagation of a scalar wave function in a stochastic medium was considered. An obvious generalization of this is the solution of the stochastic vector wave equation. The statistical measure of interest in this case is the coherency matrix (Born and Wolf, 1964). The coherency matrix with elements in a cartesian coordinate system

$$\underline{J} = \begin{bmatrix} \langle \bar{E}_x \bar{E}_x^* \rangle & \langle \bar{E}_x \bar{E}_y^* \rangle \\ \langle E_y E_x^* \rangle & \langle E_y E_y^* \rangle \end{bmatrix} \tag{5.1}$$

can be used to study the change of polarization of a wave as it propagates through a stochastic medium. It is conjectured that the Green's dyadics (Levine and Schwinger, 1951; and Van Bladel, 1964) can be used to find the statistical measures of the stochastic vector wave equation. The mathematical manipulations with the Green's dyadics are more complicated than the use of scalar Green's functions, but the general procedure for construction of stochastic Green's functions (or possibly stochastic Green's dyads) would be analogous to the scalar case. This problem would constitute a generalization of the stochastic scalar wave equation solution to the stochastic vector case.

BIBLIOGRAPHY

1. Adomian, G., "Linear Stochastic Operators," Ph.D. thesis, Department of Physics, University of California, Los Angeles (1961).

2. _____, "Linear Stochastic Operators," Rev. Modern Phys. 35, No. 1, pp. 185-207 (1963).

3. _____, "Stochastic Green's Functions," Annual Symposia in Applied Math., Vol. XVI, American Math. Soc., pp. 1-39 (1964).

4. _____, "Theory of Random Systems," Trans. of the 4th Prague Conference on Information Theory, Statistical Decision Functions, Random Processes, Publ. House Czechoslavak Acad. Sci., Prague, pp. 205-222 (1967).

5. _____, "Closure Approximations in Hierarchy Equations," a paper presented at American Physical Society Spring Meeting in Washington, D. C. (1968).

6. Anderson, M. W., Jr., "Stochastic Integral Equations," Ph.D. thesis in Mathematics, Graduate Council of the University of Tennessee (1966).

7. Åström, K. J., "On a First-Order Stochastic Differential Equation," Int. Journal of Control, Vol. 1., No. 4, pp. 301-326 (1965).

8. Barrick, D. E., "A More Exact Theory of Backscattering from Statistically Rough Surfaces," Ph.D. thesis, Department of Electrical Engineering, Ohio State University (1965).

9. Beran, M. J. and Parrent, G. B., Jr., Theory of Partial Coherence, Prentice-Hall, Englewood Cliffs, N. J. (1964).

10. Bharucha-Reid, A. T., "On Random Solutions of Fredholm Integral Equations," Amer. Math. Soc. Bulletin, Vol. 66, pp. 104-109 (1960).

11. _____, "On the Theory of Random Equations," Annual Symposia in Applied Math., Vol. XVI, Amer. Math. Soc., pp. 40-69 (1964).

12. _____, "On Random Solutions of Integral Equations in Banach Spaces," Trans. of the 2nd Prague Conference on Information Theory, Statistical Decision Functions, Random Processes, Publ. House Czechoslavak Acad. Sci., Prague, pp. 27-48 (1965).

13. Blanc-Lapierre, A., and Fortet, R., Theory of Random Functions, translated by Gani, J., Gordon and Breach, New York (1965).

14. Booker, H. B., and Gordon, W. E., "A Theory of Radio Scattering in the Troposphere," Proc. of IRE, Vol. 38, pp. 401-412 (1950).

15. _____, "Radio Scattering in the Lower Ionosphere," J. of Geophysical Research, Vol. 64, No. 12, pp. 2164-2177 (1959).

16. Born, M., and Wolf, E., Principles of Optics, 2nd revised edition, Pergamon Press, Oxford, Chapter X (1964).

17. Brown, W. P., Jr., "Propagation in Random Media - Cumulative Effect of Weak Inhomogeneities," IEEE Trans. on Antennas and Propagation, Vol. AP-15, pp. 81-89 (1967).

18. Bugnolo, D. S., "Correlation Function and Power Spectra of Radio Links Affected by Random Dielectric Noise," IRE Trans. on Antennas and Propagation, Vol. AP-7, pp. 137-141 (1959).

19. _____, "Radio Star Scintillation and Multiple Scattering in the Ionosphere," IRE Trans. on Antennas and Propagation, Vol. AP-9, pp. 89-96 (1961).

20. Caughey T. K., and Dienes, J. R., "The Behavior of Linear Systems with Random Parametric Excitation," J. Math and Phys., Vol. 41, pp. 300-318 (1962).

21. Chelpanov, B., "Vibration of a Second-order System with a Randomly Varying Parameter," Appl. Math. Mech. 26, No. 4 (1962).

22. Chernov, L. A., Wave Propagation in Random Medium, McGraw-Hill, New York (1960).

23. Clay, C. S., "Fluctuations of Sound Reflected from the Sea Surface," J. Acoust. Soc. Am., Vol. 32, No. 12, pp. 1547-1551 (1960).

24. Comstock, C., "On the Autocorrelation of Random Inhomogeneities," J. Acoust. Soc. Am., Vol. 36, No. 8, pp. 1534-1536 (1964).

25. _____, "On Wave Propagation in a Random Medium," J. of Geophysical Research, Vol. 70, No. 9, pp. 2263-2264 (1965).

26. Courant, R. and Hilbert, D., Methods of Mathematical Physics, Interscience, New York (1953).

27. Cramér, H., and Leadbetter, M. R., Stationary and Related Stochastic Processes, Sample Function Properties and Their Applications, John Wiley, New York (1967).

28. Davenport, W. B., Jr., and Root, W. L., An Introduction to the Theory of Random Signals and Noise, McGraw-Hill, New York (1958).

29. DeRusso, P. M., Roy, R. J., and Close, C. M., State Variables for Engineers, John Wiley, New York (1965).

30. Deutsch, R., Nonlinear Transformations of Random Processes, Prentice-Hall, Englewood Cliffs, New Jersey (1962).

31. Doob, J. L., Stochastic Processes, John Wiley, New York (1953).

32. Eckart, C., "The Scattering of Sound from the Sea Surface," J. Acoust. Soc. Am., Vol. 25, No. 3, pp. 566-570 (1953).

33. Edwards, D. A., and Moyal, J. E., "Stochastic Differential Equations," Proc. Cambridge Philos. Soc., Vol. 51, pp. 663-677 (1955).

34. Gibson, W. E., "An Exact Solution for a Class of Stochastic Partial Differential Equations," SIAM J. Appl. Math., Vol. 15, No. 5 (1967).

35. Goldstein, J. A., "Stochastic Differential Equations and Non-Linear Semi-groups," Ph.D. thesis in Mathematics, Carnegie Institute of Technology (1967).

36. Hanš, O., "Random Operator Equations," Proc. 4th Berkeley Sympos. on Math. Statist. and Prob., Vol. II, University of California Press, Berkeley California, pp. 185-202 (1961).

37. Hildebrand, F. B., Methods of Applied Mathemat _, Prentice-Hall, New York (1952).

38. Hoffman, W. C., "The Electromagnetic Field in a Randomly Inhomogeneous Medium," IRE Trans. on Antennas and Propagation, Vol. AP-7, pp. S301-S306 (1959).

39. _____, "Wave Propagation in a General Random Continuous Medium," Annual Symposia in Applied Math., Vol. XVI, Amer. Math. Soc., pp. 117-144 (1964).

40. Jacobson, A. D., "An Analysis of the Second Moment of Fluctuating Electromagnetic Fields, Part I: Theory, Part II: An Application to Measurement," IEEE Trans. on Antennas and Propagation, Vol. AP-15, pp. 24-40 (1967).

41. Kampé de Fériet, J., "Random Integrals of Differential Equations," in Lectures on Modern Mathematics, Saaty, T. L., ed., Vol. III, John Wiley, New York (1965).

42. Kantorovich, L. V. and Krylov, V. I. Approximate Methods of Higher Analysis, translated by Benster, C. D., Interscience, New York (1958).

43. Keller, J. B., "Stochastic Equations and Wave Propagation in Random Media," Annual Symposia in Applied Math., Vol. XVI, Amer. Math. Soc., pp. 145-170 (1964).

44. Khintchine, A., "Korrelation-theorie der stationare stochastischen Prozesse," Math. Ann., Vol. 109, pp. 604-615 (1934).

45. Kraichnan, R. H., "The Closure Problem of Turbulence Theory," Annual Symposia in Applied Math., Vol. XIII, Amer. Math. Soc., pp. 199-225 (1962).

46. Laning, J. H., Jr., and Battin, R. H., Random Processes in Automatic Control, McGraw-Hill, New York (1956).

47. Lapidus, L., and Luus, R., Optimal Control of Engineering Processes, Blaisdell, Waltham, Mass. (1967).

48. Lax, M., "Multiple Scattering of Waves," Rev. Modern Phys., Vol. 23, No. 4, pp. 287-310 (1951).

49. Lefschetz, S., "Geometric Differential Equations: Recent Past and Proximate Future," in Differential Equations and Dynamical Systems, Hale, J. K., and LaSalle, J. P., editors, Academic Press, New York (1967).

50. Levine, H. and Schwinger, J., "On the Theory of Electromagnetic Wave Defraction by an Aperture in an Infinite Plane Conducting Screen," Theory of Electromagnetic Waves, Symposium, Interscience, New York (1951).

51. Loève, M., Probability Theory, Van Nostrand, Princeton (1963).

52. Mandel, L., and Wolf, E., "Coherence Properties of Optical Fields," Rev. of Modern Physics, Vol. 37, No. 2, pp. 231-287 (1965).

53. Middleton, D., An Introduction to Statistical Communication Theory, McGraw-Hill, New York (1960).

54. Miller, K. S., Multidimensional Gaussian Distributions, John Wiley, New York (1964).

55. Mintzer, D., "Wave Propagation in a Randomly Inhomogeneous Medium I," J. Acoust. Soc. Am., Vol. 25, No. 5, pp. 922-927 (1953).

56. _____, "Wave Propagation in a Randomly Inhomogeneous Medium II," J. Acoust. Soc. Am., Vol. 25, No. 6, pp. 1107-1111 (1953).

57. _____, "Wave Propagation in a Randomly Inhomogeneous Medium III," J. Acoust. Soc. Am., Vol. 26, No. 2, pp. 186-190 (1954).

58. Moyal, J. E., "Stochastic Processes and Statistical Physics," J. Roy Statist. Soc. Ser. B, Vol. II, pp. 150-210 (1949).

59. O'Neill, E. L., Introduction to Statistical Optics, Addison Wesley, Reading, Mass. (1963).

60. Papoulis, A., Probability, Random Variables, and Stochastic Processes, McGraw-Hill, New York (1965).

61. Pugachev, V. S., Theory of Random Functions and its Application to Control Problems, translated by Blunn, O. M., and Johnson, N. L., Pergamon Press, Oxford (1965).

62. Richardson, J. M., "The Application of Truncated Hierchy Techniques in the Solution of a Stochastic Linear Differential Equation," Annual Symposia in Applied Math., Vol. XVI, Amer. Math. Soc., pp. 290-302 (1964).

63. Rosenbloom, A., "Analysis of Randomly Time-varying Linear Systems," Ph.D. thesis, University of California, Los Angeles, California (1954).

64. Samuels, J. C., and Eringen, A. C., "On Stochastic Linear Systems," J. Math. and Phys. 38, pp. 83-103 (1959).

65. Schultz, D. G., and Melsa, J. L., State Functions and Linear Control Systems, McGraw-Hill, New York (1967).

66. Špaček, A., "Regularity Properties of Random Transforms," Czechoslovak Math. J., 5(80), pp. 143-151 (1955).

67. _____, "Zufällige Gleichungen," Czechoslovak Math. J., 5(80), pp. 462-466 (1955).

68. Special Issue on Partial Coherence, IEEE Trans. on Antennas and Propagation, Vol. AP-15, No. 1 (1967).

69. Strand, J. L., "Stochastic Ordinary Differential Equations," Ph.D. thesis, Department of Mathematics, University of California, Berkeley (1967).

70. Sveshnikov, A. A., Applied Methods of the Theory of Random Functions, translated by Berry, J., and Haller, L., Pergamon Press, Oxford (1966).

71. Syski, R., "Stochastic Differential Equations," in Modern Non-Linear Equations, by Saaty, T. L., McGraw-Hill, New York (1967).

72. Tatarski, V. I., Wave Propagation in a Turbulent Medium, translated by Silverman, R. A., McGraw-Hill, New York (1961).

73. Tikhonov, V. I., "Fluctuation Action in the Simplest Parametric Systems," Automat. Remote Control, 19, pp. 705-711 (1958).

74. Twersky, V., "On Propagation in Random Media of Discrete Scatters," Annual Symposia in Applied Math., Vol. XVI, Amer. Math. Soc. pp. 84-116 (1964).

75. Van Bladel, J., Electromagnetic Fields, McGraw-Hill, New York (1964).

76. Wheelon, A. D., "Radio-Wave Scattering by Tropospheric Irregularities," J. of Research of the National Bureau of Standards - D., Vol. 63D, No. 2, pp. 205-233 (1959).

77. Zaanen, A. C., Linear Analysis, Interscience, New York (1953).

78. Zadeh, L. A., and Desoer, C. A., Linear System Theory, McGraw-Hill, New York (1963).